

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____
 An Unit: _____
 Phone Number 30 _____
 Serial Number: _____
 Results Format Preferred (circle): PAPER DISK E-MAIL _____
 Date: _____

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched.
 Include the elected species or structures, key words, synonyms, acronyms, and registry numbers, and combine with the concept or
 utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if
 known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____
 Inventors (please provide full names): _____
 Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the
 appropriate serial number.

STAFF USE ONLY

Searcher: W. Moore, Jr. NA Sequence (#) 2 Type of Search _____
 Searcher Phone #: 308-4501 AA Sequence (#) 2
 Searcher Location: Room 811 Structure (#) _____
 Date Searcher Picked Up: 9/5/01 Bibliographic _____
 Date Completed: 9/5/01 Litigation _____
 Searcher Prep & Review Time: 4m Fulltext _____
 Clerical Prep Time: 4m Patent Family _____
 Online Time: 4m Other _____
 PT0-1590 (1-2000)

Vendors and cost where applicable

STN _____
 Dialog _____
 Quest/Orbit _____
 Dr. Link _____
 Lexis/Nexis _____
 Sequence Systems AB5502 _____
 WWW/Internet _____
 Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:03 ; Search time 21.04 Seconds
(without alignments)
994.072 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 1848
Sequence: 1 MLLGLLLTSLAQRTGT.....DVALEHHERCDVCVCGNAGG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1848	100.0	345	21	Mouse zveg3, SEQ
2	1848	100.0	345	21	Murine vascular en
3	1848	100.0	345	21	A murine platelet-
4	1667	90.2	345	20	Human VEGF-E prote
5	1667	90.2	345	20	Human PRO200 prote
6	1667	90.2	345	20	Human vascular end
7	1667	90.2	345	21	Human zveg3, SEQ
8	1667	90.2	345	21	Human platelet-der
9	1667	90.2	345	21	Human PRO200 (UNQ1
10	1667	90.2	345	21	Human RACE generat
11	1667	90.2	345	21	Human VEGF-X prote

12	1667	90.2	345	21	AAB10644	Human VEGF-X prote
13	1667	90.2	345	21	AAB10650	Human 990126veg p
14	1667	90.2	345	21	AAB10651	Human VEGF-X prote
15	1667	90.2	345	21	AAB19578	Human PRO200 (vasc
16	1667	90.2	345	21	AAB33414	Human PRO200 prote
17	1667	90.2	345	21	AAB24412	Human PRO713 prote
18	1667	90.2	345	21	AAB01419	Human TANGO 128.
19	1667	90.2	345	21	AAY96858	Human growth facto
20	1667	90.2	345	21	AAY59285	Bone morphogenic p
21	1667	90.2	345	22	AAB50980	Human PRO200 prote
22	1667	90.2	345	22	AAB49895	Human PRO200 prote
23	1667	90.2	345	22	AAB53074	Human angiogenesis
24	1667	90.2	374	21	AAB10639	Human VEGF-X prote
25	1664	90.0	345	21	AAY84557	Amino acid sequenc
26	1659	89.8	345	21	AAB10634	Human VEGF-X homol
27	1659	89.8	345	21	AAB10636	Human VEGF-X prote
28	1659	89.8	345	21	AAB03003	Human growth facto
29	1576	85.3	354	21	AAB10640	Human VEGF-X prote
30	1576	85.3	354	21	AAB10641	Human VEGF-X prote
31	1559	84.4	318	21	AAY84558	A fragment of plat
32	1524.5	82.5	339	21	AAB58438	Lung cancer associ
33	1267.5	68.6	282	21	AAB10653	Human VEGF-X prote
34	1145	62.0	297	21	AAY59286	Bone morphogenic p
35	852	46.1	180	21	AAB01427	Murine TANGO 128.
36	752	40.7	370	21	AAB48653	Human growth facto
37	752	40.7	370	21	AAY98664	SEQ. ID. 37 from W
38	752	40.7	370	21	AAY71130	Human Platelet Der
39	752	40.7	370	22	AAB60888	Human VEGF-G prote
40	750.5	40.6	322	21	AAY71129	Human Platelet Der
41	746.5	40.4	370	21	AAB48663	Mouse growth facto
42	744.5	40.3	370	22	AAB60895	Human VEGF-G prote
43	720	39.0	167	21	AAB10652	Human VEGF-X prote
44	705	38.1	168	21	AAB10643	Human VEGF-X CUB-1
45	610	33.0	149	21	AAB10642	Human VEGF-X PDGF-

ALIGNMENTS

RESULT 1

AAB48658
 ID AAB48658 standard; Protein; 345 AA.
 XX AAB48658;
 AC AAB48658;
 XX
 DT 09-MAR-2001 (first entry)
 DT
 XX
 DE Mouse zveg3, SEQ ID NO:35.
 XX
 KW Mouse; zveg3; zveg3 fusion; growth factor homologue; VEGF/PDGF family;
 KW murine; CUB domain; PDGF-like activity; mitogenic; osteogenic;
 KW neovascularisation; tissue repair; proliferation; differentiation;
 KW liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
 KW periodontal disease; bone fracture; wound healing; vulnery; ischaemia;
 KW immunomodulation; hepatic.
 XX
 OS Mus musculus.
 XX
 PN WO200066736-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 03-MAY-2000; 2000WO-US40047.
 XX
 PR 03-MAY-1999; 99US-0304216.
 PR 10-NOV-1999; 99US-0164463.
 PR 04-FEB-2000; 2000US-0180169.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;
 XX
 DR WPI; 2000-687541/67.

DR N-PSDB; AAC81583.
XX Growth factor homologs and the nucleic acids that encode them, useful
PT e.g. for treating liver damage, ischemia, multiple sclerosis and
PT Alzheimer's disease -
XX
XX Disclosure: Page 130-131; 143pp; English.
XX
XX The invention relates to the human growth factor homologue zvegfg4
CC (AAB48653), and nucleic acids encoding it (AAC81555). zvegfg4 is a member
CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
CC growth factor) family. zvegfg4 has a growth factor domain (AAB48654)
CC characterised by a PDGF cysteine knot structure, and a CUB domain
CC (AAB48655) which has a beta barrel structure. zvegfg4 has PDGF-like
CC activity, having mitogenic activity on fibroblasts, vascular smooth
CC muscle cells and pericytes, and has also been shown to stimulate bone
CC growth. The invention also relates to fusion proteins comprising human
CC zvegfg4 or fragments thereof, particularly human zvegfg4/human zvegfg3
CC fusions; expression constructs and host cells comprising human zvegfg4
CC nucleic acids; the recombinant expression of human zvegfg4; an antibody
CC which binds to human zvegfg4 or a fragment thereof; a method of activating
CC a cell-surface PDGF receptor using a zvegfg4-derived polypeptide; a
CC method of modulating the proliferation, differentiation, migration or
CC metabolism of bone cells, comprising exposing bone cells to
CC zvegfg4-derived polypeptides; and a method of detecting a genetic
CC abnormality in the zvegfg4 gene of a patient. zvegfg4 proteins and derived
CC fragments may be used to stimulate tissue development or repair, or
CC cellular differentiation or proliferation. They are particularly used for
CC the treatment or repair of liver damage, and may also be used to
CC modulate neurite growth (e.g., in the treatment of Alzheimer's disease or
CC multiple sclerosis). Due to their osteogenic activity, they may be used
CC in the treatment of periodontal disease and fractures. They may also be
CC used to enhance expansion and mobilisation of haematopoietic stem cells
CC and endothelial precursor stem cells, which may be useful in the
CC treatment of ischaemia, in wound healing, and in the modulation of the
CC immune system. The present sequence represents mouse zvegfg3.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1848; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.3e-182;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLLLTSALAGQRTGRTAESNLSSKLQLSDDKEQNGVODPRHVRVVTISGNGSIHS 60
DB 1 mllgl'llltsaleqqr'tgtraesnl'ssklq'lsddk'eqngv'odp'r'hv'r'vvt'isgng'sihs 60
QY 61 PKFPHYPRNMVLVRLVAVDENVRILQTFDERFGLDEPDDICKYDFVEVEEPSDGSYL 120
DB 61 pkfphtyprnmvlvrlvavdenvrilqtfderf'gldepeddick'ydfveveepsdgsyl 120
QY 121 GRWCGSGTVPGKQTSKGNHRIIRFVSDEYFSPSEFCFCHYSIIMPQVTTTSPSVLPSS 180
DB 121 grwcgsgtvpgkqtskgnhriirfvsdey'fspse'fcfch'ysiimp'qvttt'spsvlp'ss 180
QY 181 LSLDLLNNAVTAFTLEELIRYLEPDRQWDLDSLYKPTWQLLGRAFLYGRKSKVYNLNL 240
DB 181 lsldllnnavtaftleelirylepdrq'wdl'dslyk'ptw'ql'lg'rafl'ygrk'skvyn'lnl 240
QY 241 LKEEVKLYSCTPRNFVSIRIELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPRK 300
DB 241 lkeevklysc'tprn'fvsir'ielk'rtdt'ifw'pgc'llv'kr'cgnc'accl'hncn'ccq'vprk 300
QY 301 VTKKYHEVLQIRPKTVGKGLHSLTDVALEHHEECDCVCRGNAGG 345
DB 301 vt'kkyhe'vlq'irp'ktvg'kgl'hsl'tdva'lehh'ee'cd'cv'cr'nagg 345

RESULT 2

AAY96861

ID AAY96861 standard; Protein; 345 AA.

XX

AC AAY96861;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine vascular endothelial growth factor homologue, ZVEGF3.
XX
KW Vascular endothelial growth factor; homologue; zvegfg3; CUB domain;
KW Cysteine knot; platelet-derived growth factor; PDGF; neuropilin;
KW chromosome 4q28.3; cytostatic; anti-psoriatic; anti-inflammatory;
KW anti-diabetic; ophthalmological; anti-rheumatic; anti-arthritis;
KW vulnery.
OS Mus musculus.
PN WO200034474-A2.
XX
PD 15-JUN-2000.
XX
XX 07-DEC-1999; 99WO-US28968.
PF
XX 07-DEC-1998; 98US-0207120.
PR
XX 06-JUL-1999; 99US-0142576.
PR
XX 21-OCT-1999; 99US-0161653.
PR
XX 12-NOV-1999; 99US-0165255.
PR
XX (ZYMO) ZYMOGENETICS INC.
PA
XX Gao Z, Hart CE, Piddington CS, Sheppard PO, Shoemaker KE;
PI Gilbertson DG, West JW;
XX
XX WPI: 2000-423420/36.
DR N-PSDB; AAA51527.
XX
XX Novel zvegfg3 polypeptides and nucleotides encoding them useful for
PT stimulating growth of smooth muscle cells and fibroblasts comprising an
PT epitope bearing portion of a specific amino acid sequence
XX
XX Claim 1; Page 169-170; 173pp; English.
XX
XX This shows a murine ZVEGF3 a novel vascular endothelial growth factor
CC homologue. Polypeptides comprising an epitope-bearing portion human or
CC murine ZVEGF3 are claimed. The growth factors comprise a growth factor
CC domain and a CUB domain (generic sequence motifs are shown in AAY96859
CC and AAY96860). The growth factor domain is characterized by an
CC arrangement of cysteine residues and beta-strands that is characteristic
CC of the "cysteine knot" structure of the platelet-derived growth factor
CC (PDGF) family. The CUB domain shows homology to CUB domains in
CC neuropilins, human bone morphogenetic protein-1, porcine seminal plasma
CC protein, bovine acidic seminal fluid protein and Xenopus laevis
CC toll-like protein. Structural analysis and homology predict that
CC ZVEGF3 polypeptides complex with a second polypeptide to form multimeric
CC proteins. The human zvegfg3 gene has been mapped to chromosome 4q28.3.
CC ZVEGF3 is useful for stimulating the growth of fibroblasts or smooth
CC muscle cells, for activating cell surface PDGF-alpha receptor and for
CC inhibiting PDGF-alpha receptor mediated cellular processes. ZVEGF3 is
CC useful for regulating (post-development) organ growth, regeneration and
CC maintenance, as well as tissue maintenance and repair processes. ZVEGF3
CC antagonists are useful for treating cancer, rheumatoid arthritis,
CC diabetic retinopathy, ischemic limb disease, peripheral vascular
CC disease, myocardial ischemia, vascular intimal hyperplasia,
CC atherosclerosis, wound healing, chronic liver disease and haemangioma
CC formation. ZVEGF3 can also be used to modulate neurite growth and
CC development of the nervous system, and for treating neurodegenerative
CC diseases.
XX
XX Sequence 345 AA;.

Query Match 100.0%; Score 1848; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.3e-182;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLLLTSALAGQRTGRTAESNLSSKLQLSDDKEQNGVODPRHVRVVTISGNGSIHS 60

Db 1 mlllglllltsalagqrtrgtraesnlssklqlssdkqngvqdrhervvtisngsihs 60
Qy 61 PKFHTYPRNMVLRVAVDENVRIOITFDERFGLDEPDDICKYDFVEVEEPSGVSVL 120
Db 61 pkfhtyprnmvlrvlavdenvrilqitfderfgledpeddickdyfveveepsdgsvl 120
Qy 121 GRWCGSGTVPGKOTSGKNHIRIRFVSDYFSEPGFCIHYSIIMPQVTTTSPSVLPSS 180
Db 121 grwcgsgtvpgktskgnhrrfsvdysfsepgfcihysimpqvtettspsvlpss 180
Qy 181 LSLLNNAVTAFSTLEELIRYLEPDRWQVDLSLYKPTWQLLGKAFLYGKSKVNNLNL 240
Db 181 lsdlldnnavtafstleelirylepdrwqvdlslykptwqllgkaflygkkskvvnnlnl 240
Qy 241 LKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPK 300
Db 241 lkeevklyscprnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqcvpk 300
Qy 301 VTKYHEVLQLRPKTGVKGLHKSLLTDVALEHHEEDCVCRCGNAGG 345
Db 301 vtkeyhevlqlrpkgtgvkglhkslldvalehheecdvcrcgnagg 345

RESULT 3
AA584559
ID AAY84559 standard; Protein; 345 AA.
XX AC AAY84559;
XX 25-JUL-2000 (first entry)
DE A murine platelet-derived growth factor C (PDGF-C).
KW Platelet-derived growth factor C; PDGF-C; cell proliferation;
KW growth factor; heparin; connective tissue; wound healing; VEGF-F;
KW fibroblast mitogenesis; PDGF alpha receptor activation; tumour growth;
KW choriocarcinoma; Wilms tumour; megakaryoblastic leukaemia;
KW lung carcinoma; erythroleukemia; tissue remodelling.
XX Mus sp.
XX WO200018212-A2.
XX 06-APR-2000.
XX 30-SEP-1999; 99WO-US22668.
XX 30-SEP-1998; 98US-0102461.
XX 12-NOV-1998; 98US-0108109.
XX 03-DEC-1998; 98US-0110749.
XX 18-DEC-1998; 98US-0113002.
XX 21-MAY-1999; 99US-0135426.
XX 15-JUL-1999; 99US-0144022.
XX (LUDW-) *LUDWIG INST CANCER RES.
XX (UYHE-) UNIV HELSINKI LICENSING LTD.
XX Eriksson U, Aase K, Lee X, Ponten A, Uutela M, Alitalo K;
XX Oestman A, Heldin C, Betsholtz C;
XX WPI; 2000-292954/25.
XX N-PSDB; AAA12525.
XX Novel DNA encoding PDGF-C useful to stimulate or enhance proliferation,
XX differentiation, growth and motility of cells expressing the PDGF-C
XX receptor -
XX Claim 27; Fig 6; 135pp; English.
XX The present sequence represents murine platelet-derived growth factor C
XX (PDGF-C) (formally designated VEGF-F). PDGF-C polypeptides have the
XX ability to stimulate and enhance proliferation or differentiation,

CC and/or growth or motility of cells expressing a PDGF-C receptor.
CC PDGF-C polypeptides can be used in pharmaceuticals for promoting cell
CC proliferation, preferably in combination with one other growth factor
CC and heparin. Pharmaceuticals comprising PDGF-C polypeptides can also
CC be used for stimulating connective tissue or wound healing. The
CC PDGF-C polypeptide can be enzymatically processed to generate the active
CC truncated form of PDGF-C and used to regulate the receptor-binding
CC specificity of PDGF-C. PDGF-C can also be used to promote fibroblast
CC mitogenesis in a mammal and to induce PDGF alpha receptor activation.
CC PDGF-C antagonists can be used to inhibit tumour growth of a tumour
CC expressing PDGF-C in a mammal. Specific types of human tumours, e.g.
CC choriocarcinoma, Wilms tumour, megakaryoblastic leukaemia, lung carcinoma
CC and erythroleukemia, can be identified by testing for expression of
CC PDGF-C. PDGF-C antagonists can also be used to inhibit tissue
CC remodelling during invasion of tumour cells into a normal population of
CC cells. Antagonists can also be used to treat fibrotic conditions,
CC especially found in the lung, kidney or liver.
XX Sequence 345 AA;
SQ

Query Match 100.0%; Score 1848; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.3e-182;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLGLLLLSALAGQRTGTRAESNLSSKLQLSSDKQNGVQDRHERVVTISNGSIHS 60
Db 1 mlllglllltsalagqrtrgtraesnlssklqlssdkqngvqdrhervvtisngsihs 60
Qy 61 PKFHTYPRNMVLRVAVDENVRIOITFDERFGLDEPDDICKYDFVEVEEPSGVSVL 120
Db 61 pkfhtyprnmvlrvlavdenvrilqitfderfgledpeddickdyfveveepsdgsvl 120
Qy 121 GRWCGSGTVPGKOTSGKNHIRIRFVSDYFSEPGFCIHYSIIMPQVTTTSPSVLPSS 180
Db 121 grwcgsgtvpgktskgnhrrfsvdysfsepgfcihysimpqvtettspsvlpss 180
Qy 181 LSLLNNAVTAFSTLEELIRYLEPDRWQVDLSLYKPTWQLLGKAFLYGKSKVNNLNL 240
Db 181 lsdlldnnavtafstleelirylepdrwqvdlslykptwqllgkaflygkkskvvnnlnl 240
Qy 241 LKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPK 300
Db 241 lkeevklyscprnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqcvpk 300
Qy 301 VTKYHEVLQLRPKTGVKGLHKSLLTDVALEHHEEDCVCRCGNAGG 345
Db 301 vtkeyhevlqlrpkgtgvkglhkslldvalehheecdvcrcgnagg 345

RESULT 4
AA533679
ID AAY33679 standard; Protein; 345 AA.
XX AC AAY33679;
XX 11-JAN-2000 (first entry)
XX Human VEGF-E protein.
XX VEGF-E; human; vascular endothelial cell growth factor; wound repair;
XX treatment; cardiovascular disorder; endothelial disorder; therapy;
XX tissue generation; regeneration; cardiac hypertrophy; cancer; detection;
XX angiogenic disorder; age-related macular degeneration; vascular disease;
XX neovascularization; tumor; gene mapping.
XX Homo sapiens.
XX WO9947677-A2.
XX 23-SEP-1999.
XX 10-MAR-1999; 99WO-US05190.

XX 17-MAR-1998; 98US-0040220.
 PR 02-NOV-1998; 98US-0184216.
 XX (GETH) GENENTECH INC.
 XX Ferrara N, Kuo SS;
 XX WPI: 1999-580306/49.
 DR N-PSDB: AAZ23691.
 XX
 PR New growth factor polypeptide useful for treating cardiovascular or
 PT endothelial disorders, e.g. cardiac hypertrophy
 XX Claim 1; Fig 2; 12pp; English.
 XX
 CC This invention describes the isolation of a novel human vascular
 CC endothelial cell growth factor-E (VEGF-E) polypeptide which has
 CC tranquilizer, vulnery and cardiant activity. VEGF-E can be administered
 CC therapeutically, especially by expressing encoding polynucleotides, to
 CC treat cardiovascular or endothelial disorders in mammals, especially
 CC humans. It is useful in wound repair and tissue generation and
 CC regeneration, and may especially be used to treat cardiac hypertrophy
 CC It can be combined with a carrier in pharmaceutical compositions, which
 CC can be administered to treat disorders as above. VEGF-E can be used to
 CC screen for antagonists and agonists, and the antagonists administered to
 CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
 CC age-related macular degeneration. It can be used to generate antibodies,
 CC useful therapeutically as antagonists, as above. The antibodies are also
 CC useful to detect VEGF-E polypeptide, especially to diagnose
 CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
 CC vascular disease, or neovascularization associated with tumor formation),
 CC by contacting the antibody with a tissue sample and detecting formation
 CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
 CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
 CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
 CC tissue samples. They can also be used to diagnose a disease or
 CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
 CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
 CC detecting a mutation in the VEGF-E encoding sequence isolated from a
 CC sample. They may also be used to produce probes useful to detect related
 CC sequences or for gene mapping. This sequence represents the human VEGF-E
 CC protein described in the method of the invention.
 XX
 SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 20; Length 345;
 Best Local Similarity 87.0%; Pred. No. 1.2e-163;
 Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLSALAGORTGTAEISNLSSKLQSSDKQNGVDPRHVVITISNGSIHS 60
 Db 1 mslflllltsalagqrqgtgaesnlsskfgfnsknegvqdpqherlitystngsihs 60

QY 61 PKFPHYPRNNVLVRLVAVDENVRQLTFDERFLEDPEDDICKYDFVEVEEPSDGSVL 120
 Db 61 prfphtyprntvlrvlvaeevwwqltfderrfledpeddickdyfveveepsdgtil 120

QY 121 GRWCGSTVPGKQTSKGNHIRVRSDYFPPSEPGFCFTHYSIMPOVTTTSPSVLPSPS 180
 Db 121 grwcgsgtvpgkqtskgnqirrvsvdeyfpsepgfcfthynivmpqfteavspsvlpssa 180

QY 181 LSLDLNNVATFSTLEIRYLPDRQVLDLSLYKWTQLLGLKFLYKSKVNLNL 240
 Db 181 lpldlannatfstlelrylperwqlldlrylrvqlwllgkafgrksrvvdlal 240

QY 241 LKEEVKLYSCTPRNFSVIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCVPRK 300
 Db 241 lteevrlyscprnfsvireelkrttdtlfwpgcllvkrccgncacclhncnccvcprk 300

QY 301 VTKKYHEVLQLRPKTGKGLHKSITDVALEHHEECDCVCRGNAGG 345
 Db 301 vtkkyhevlqlrpktdgkghksitdvalhheecdcvcrngagg 345

Db 301 vtkkyhevlqlrpktdgkghksitdvalhheecdcvcrngagg 345

RESULT 5
 AAY41766
 ID AAY41766 standard; Protein; 345 AA.
 XX
 AC AAY41766;
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO200 protein sequence.
 XX
 KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KW secreted protein; transmembrane protein.
 OS Homo sapiens.
 XX
 PN WO9946281-A2.
 XX
 PD 16-SEP-1999.
 XX
 PF 08-MAR-1999; 99WO-US05028.
 XX
 PR 10-MAR-1998; 98US-0077450.
 PR 11-MAR-1998; 98US-0077632.
 PR 11-MAR-1998; 98US-0077641.
 PR 11-MAR-1998; 98US-0077649.
 PR 12-MAR-1998; 98US-0077791.
 PR 13-MAR-1998; 98US-0078004.
 PR 17-MAR-1998; 98US-0040220.
 PR 20-MAR-1998; 98US-0078886.
 PR 20-MAR-1998; 98US-0078910.
 PR 20-MAR-1998; 98US-0078936.
 PR 20-MAR-1998; 98US-0078939.
 PR 25-MAR-1998; 98US-0079294.
 PR 26-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079663.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 27-MAR-1998; 98US-0079728.
 PR 27-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 30-MAR-1998; 98US-0079923.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 31-MAR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 01-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081070.
 PR 08-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 09-APR-1998; 98US-0081229.
 PR 15-APR-1998; 98US-0081817.
 PR 15-APR-1998; 98US-0081838.
 PR 15-APR-1998; 98US-0081952.
 PR 15-APR-1998; 98US-0081955.
 PR 21-APR-1998; 98US-0082568.
 PR 21-APR-1998; 98US-0082569.
 PR 22-APR-1998; 98US-0082700.
 PR 22-APR-1998; 98US-0082704.
 PR 22-APR-1998; 98US-0082804.
 PR 23-APR-1998; 98US-0082767.
 PR 23-APR-1998; 98US-0082796.
 PR 27-APR-1998; 98US-0083336.
 PR 28-APR-1998; 98US-0083322.
 PR 29-APR-1998; 98US-0083392.

PR 29-APR-1998; 98US-0083495.
 PR 29-APR-1998; 98US-0083496.
 PR 29-APR-1998; 98US-0083499.
 PR 29-APR-1998; 98US-0083500.
 PR 29-APR-1998; 98US-0083545.
 PR 29-APR-1998; 98US-0083554.
 PR 29-APR-1998; 98US-0083558.
 PR 29-APR-1998; 98US-0083559.
 PR 30-APR-1998; 98US-0083742.
 PR 05-MAY-1998; 98US-0084366.
 PR 06-MAY-1998; 98US-0084411.
 PR 07-MAY-1998; 98US-0084598.
 PR 07-MAY-1998; 98US-0084600.
 PR 07-MAY-1998; 98US-0084627.
 PR 07-MAY-1998; 98US-0084637.
 PR 07-MAY-1998; 98US-0084639.
 PR 07-MAY-1998; 98US-0084640.
 PR 07-MAY-1998; 98US-0084643.
 PR 13-MAY-1998; 98US-0085323.
 PR 13-MAY-1998; 98US-0085338.
 PR 13-MAY-1998; 98US-0085339.
 PR 15-MAY-1998; 98US-0085573.
 PR 15-MAY-1998; 98US-0085579.
 PR 15-MAY-1998; 98US-0085580.
 PR 15-MAY-1998; 98US-0085582.
 PR 15-MAY-1998; 98US-0085689.
 PR 15-MAY-1998; 98US-0085697.
 PR 15-MAY-1998; 98US-0085700.
 PR 15-MAY-1998; 98US-0085704.
 PR 18-MAY-1998; 98US-0086023.
 PR 22-MAY-1998; 98US-0086392.
 PR 22-MAY-1998; 98US-0086414.
 PR 22-MAY-1998; 98US-0086430.
 PR 22-MAY-1998; 98US-0086486.
 PR 28-MAY-1998; 98US-0087098.
 PR 28-MAY-1998; 98US-0087106.
 PR 28-MAY-1998; 98US-0087208.
 PR 30-JUL-1998; 98US-0094651.
 PR 11-SEP-1998; 98US-0100038.
 XX (GETH) GENENTECH INC.

XX PA-
 XX PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX XX

DR WPI; 1999-551358/46.
 DR N-PSDB; AAZ34296.

XX New secreted and transmembrane polypeptides and their polynucleotides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders

XX Claim 12; Fig 207; 530pp; English.

XX The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to create transgenic
 CC animals. The proteins can be used to treat a variety of diseases and
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. AAZ33891 to
 CC AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
 CC polypeptide sequence given in the exemplification of the present
 CC invention.

XX Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 20; Length 345;
 Best Local Similarity 87.0%; Pred. No. 1.2e-163;
 Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLTALAGORTGTRAESNLSSKQLQSSDKEQGVQDPHERVVTISNGSIHS 60
 Db 1 mslfglllltsalagrggtqaesnlsskfqfnsnkedqgvdpqheriitvtngsihs 60
 QY 61 PKPHTYPRNMVLRVAVDENVRIOQLTDERFGLDEPDICKYDFVEVEEPSDGSVL 120
 Db 61 prfptyprntvrvlravdeenvriqltderfgleddedickdydveveepsdgtiil 120
 QY 121 GRWCGSGTVFGKQTSKGNHIRFVSDYFSPFSGFCIHYSIIMPQVTTTSPVLPSS 180
 Db 121 grwcgsgtvpgkqskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180
 QY 181 LSLDLLNNAVTAFSTLEELIRYLEPDRQWQVLDLSLYKPTWOLLKAFLYGKKSVNLL 240
 Db 181 lpldllnnaitafstledliryleperwqldledlyrptwllgkafvgrksrvdlnl 240
 QY 241 LKEEVKLYSCTPRNFSYSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCQCQVPRK 300
 Db 241 lteevrlyscprnfsysireelkrttdtiffpgcllvkrccgncacclhncqcqvpsk 300
 QY 301 VTKKYHEVLQRLPKTGVKGLKSLTDVALEHHEEDCVCRCGNAGG 345
 Db 301 vtkkyhevlqlrpkrtgvrghlksltovalehheecdvcrcgstgg 345
 RESULT 6
 AAY30023
 ID AAY30023 standard; Protein; 345 AA.
 XX
 AC AAY30023;
 XX DT 11-OCT-1999 (first entry)
 XX DE Human vascular endothelial growth factor related protein.
 XX KW Vascular endothelial growth factor related protein; VEGF-R protein;
 KW tissue growth inhibition; tumour growth; cancer; tissue growth;
 KW angiogenesis; coronary artery blockage.
 XX OS Homo sapiens.
 XX PN WO9937671-A1.
 XX PD 29-JUL-1999.
 XX PF 26-JAN-1999; 99WO-US01574.
 XX PR 31-AUG-1998; 98US-0098548.
 PR 27-JAN-1998; 98US-0072635.
 PR 05-JUN-1998; 98US-0088089.
 PR 24-JUN-1998; 98US-0090544.
 XX (ELIL) LILLY & CO ELI.
 XX PI Dou S, Na S, Song HY;
 XX WPI; 1999-458680/38.
 DR N-PSDB; AAX86352.
 XX A vascular endothelial growth factor related protein and related
 PT polynucleotide, useful for identifying antagonists and binding
 PT compounds
 XX Claim 1; Page 56-58; 62pp; English.
 XX The present sequence represents a vascular endothelial growth factor
 CC related (VEGF-R) protein. VEGF-R can be used in assays to identify
 CC compounds that bind to it or that antagonize its activity. VEGF-R
 CC antagonists (e.g. anti-VEGF-R antibodies) are useful for inhibiting
 CC tissue growth. This is useful for inhibiting tumour growth and for
 CC treating cancer. VEGF-R itself can be used to stimulate tissue
 CC growth, angiogenesis and to treat coronary artery blockage. The
 CC VEGF-R coding sequence can be used for the recombinant production of

CC the VEGF-R protein.
 SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 20; Length 345;
 Best Local Similarity 87.0%; Pred. No. 1.2e-163;
 Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLTALAGQRTGTRAESNLSSKQLQSSDKEQNGVDPDRHVRVVTISGNGSIHS 60
 DB 1 mslfglllltsalagrgtgtaesnlsskfssnkeqngvdpqphertitvtngsihs 60

QY 61 PKPPTYPRNMVWLVRVAVDENVRQLTDFERFGLDEPDDICKYDFVEVEPSPSGSVL 120
 DB 61 prphtyprntvwlrvaveenvvqltderfgleddedddickdyfvevepsdgtl 120

QY 121 GRWCGSGTVPGKQTSKGNHIRIRFVSDEYFPSEPGFCIHYSIIMQVTTTSPSVLPSS 180
 DB 121 grwcsgstvpdkqtskgnqirirfvsdeyfpsepgfcihynlvmpqfteavspvlpssa 180

QY 181 LSLDLLNNAVTAFTLEELIRYLEPDRQVLDLSLYKPTWQLLGAFLYGKSKVVNLNL 240
 DB 181 lpldllnnavtafsciedliryleperwqlldledlyrptwqlgkafrgrksrvvdl 240

QY 241 LKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
 DB 241 lteevrlyscprnfsvsireelkrttdtfwpgcllvkrccgncacclhncnccqvpk 300

QY 301 VTKKYHEVLQRLPKTGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
 DB 301 vtkeyhevlqlrpkgtgvrghlsldvalehheecdcvcrgstg 345

RESULT 7
 AAB48657
 ID AAB48657 standard; Protein: 345 AA.
 XX AC AAB48657;
 XX DT 09-MAR-2001 (first entry)
 XX DE Human zvegfg3, SEQ ID NO:33.
 XX KW Human; zvegfg3; zvegfg4 fusion; growth factor homologue; VEGF/PDGF family;
 KW CUB domain; PDGF-like activity; mitogenic; osteogenic;
 KW neovascularisation; tissue repair; proliferation; differentiation;
 KW liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
 KW periodontal disease; bone fracture; wound healing; vulnerability; ischaemia;
 KW immunomodulation; hepatic.
 XX OS Homo sapiens.
 XX PN WO200066736-A1.
 XX PD 09-NOV-2000.
 XX PF 03-MAY-2000; 2000WO-US40047.
 XX PR 03-MAY-1999; 99US-0304216.
 XX PR 10-NOV-1999; 99US-0164463.
 XX PR 04-FEB-2000; 2000US-0180169.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;
 XX WPI: 2000-687541/67.
 XX DR N-PSDB; AAC81582.
 XX Growth factor homologs and the nucleic acids that encode them, useful
 PT e.g. for treating liver damage, ischemia, multiple sclerosis and
 PT Alzheimer's disease -

Claim 48; Page 125-126; 143pp; English.
 The invention relates to the human growth factor homologue zvegfg4 (AA848653), and nucleic acids encoding it (AAC81555). zvegfg4 is a member of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial growth factor) family. zvegfg4 has a growth factor domain (AA848654) characterised by a PDGF cysteine knot structure, and a CUB domain (AA848655) which has a beta barrel structure. zvegfg4 has PDGF-like activity, having mitogenic activity on fibroblasts, vascular smooth muscle cells and pericytes, and has also been shown to stimulate bone growth. The invention also relates to fusion proteins comprising human zvegfg4 or fragments thereof, particularly human zvegfg4/human zvegfg3 fusions; expression constructs and host cells comprising human zvegfg4 nucleic acids; the recombinant expression of human zvegfg4; an antibody which binds to human zvegfg4 or a fragment thereof; a method of activating a cell-surface PDGF receptor using a zvegfg4-derived polypeptide; a method of modulating the proliferation, differentiation, migration or metabolism of bone cells, comprising exposing bone cells to zvegfg4-derived polypeptides; and a method of detecting a genetic abnormality in the zvegfg4 gene of a patient. zvegfg4 proteins and derived fragments may be used to stimulate tissue development or repair, or cellular differentiation or proliferation. They are particularly used for the treatment or repair of liver damage, and may also be used to modulate neurite growth (e.g., in the treatment of Alzheimer's disease or multiple sclerosis). Due to their osteogenic activity, they may also be used in the treatment of periodontal disease and fractures. They may also be used to enhance expansion and mobilisation of haematopoietic stem cells and endothelial precursor stem cells, which may be useful in the treatment of ischaemia, in wound healing, and in the modulation of the immune system. The present sequence represents human zvegfg3.

Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;
 Best Local Similarity 87.0%; Pred. No. 1.2e-163;
 Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLTALAGQRTGTRAESNLSSKQLQSSDKEQNGVDPDRHVRVVTISGNGSIHS 60
 DB 1 mslfglllltsalagrgtgtaesnlsskfssnkeqngvdpqphertitvtngsihs 60

QY 61 PKPPTYPRNMVWLVRVAVDENVRQLTDFERFGLDEPDDICKYDFVEVEPSPSGSVL 120
 DB 61 prphtyprntvwlrvaveenvvqltderfgleddedddickdyfvevepsdgtl 120

QY 121 GRWCGSGTVPGKQTSKGNHIRIRFVSDEYFPSEPGFCIHYSIIMQVTTTSPSVLPSS 180
 DB 121 grwcsgstvpdkqtskgnqirirfvsdeyfpsepgfcihynlvmpqfteavspvlpssa 180

QY 181 LSLDLLNNAVTAFTLEELIRYLEPDRQVLDLSLYKPTWQLLGAFLYGKSKVVNLNL 240
 DB 181 lpldllnnavtafsciedliryleperwqlldledlyrptwqlgkafrgrksrvvdl 240

QY 241 LKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
 DB 241 lteevrlyscprnfsvsireelkrttdtfwpgcllvkrccgncacclhncnccqvpk 300

QY 301 VTKKYHEVLQRLPKTGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
 DB 301 vtkeyhevlqlrpkgtgvrghlsldvalehheecdcvcrgstg 345

RESULT 8
 AAB24250
 ID AAB24250 standard; Protein: 345 AA.
 XX AC AAB24250;
 XX DT 08-FEB-2001 (first entry)
 XX DE Human platelet-derived growth factor related protein LP8.

XX Human; platelet derived growth factor related protein; LP8; VEGFh;
KW vascular endothelial growth factor h; tissue regeneration; vulnerary;
KW atherosclerosis; PDGF-related protein; antiarteriosclerotic.
XX
OS Homo sapiens.
PN WO200059940-A2.
XX 12-OCT-2000.
XX 24-MAR-2000; 2000WO-US06427.
XX 06-APR-1999; 99US-0127913.
XX (ELIL) LILLY & CO ELI.
XX Hammond LJ, Na S;
XX
XX WPI; 2000-664991/64.
DR N-PSDB; AAC64426.
XX
XX Enhancing tissue growth and promoting wound healing by administering
PT platelet-derived growth factor related protein, LP8 or its analog and
PT treating atherosclerosis by administering LP8 antagonist
XX
XX Claim 4; Page 63-64; 64pp; English.
XX
XX The present invention describes a method for enhancing tissue growth,
CC promoting wound healing or stimulating smooth muscle growth by
CC administering a platelet-derived growth factor (PDGF) related protein,
CC designated LP8 or its analogue. Also described is a method of slowing
CC the progress of atherosclerosis or treating atherosclerosis comprising
CC the administration of an LP8 antagonist. The method is useful for
CC enhancing tissue growth, promoting wound healing and stimulating smooth
CC muscle growth. Antagonists of LP8 are useful for treating
CC atherosclerosis. The present sequence represents human LP8, which is
CC also called VEGFh.
XX
SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.2e-163;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;
QY 1 MLLGLLLLSALAGORTGTRAESNLSSKLQSSDKQNGVQDPRHERVVTISGNGSIHS 60
Db 1 mslfglllltsalagqrqgqaeenlsskfqsnskeqngvqdpqherllvtstngsihs 60
QY 61 PKFPHTYPRNMVLVRLVAVDENVRIOITFDERFGLDEPDIDCKYDFVEVEEPSDGSVL 120
Db 61 prfphyprrntvlvrlvaveenwqltderfgleddedidckdyfveveepsdgtl 120
QY 121 GRWCGSTVPKOTSGKNHRIREFVDEYPPSPGFCIHYSIIMPQVTTSPVLPSS 180
Db 121 grwcgsatvpqgksgnqirirfvsdeyfpsepgfcilvnaimpqfteavspvlpsa 180
QY 181 LSLDLLNNAVATSTLEELRYLPDRWQVDSLKYPTWQLGKAFLYGKKSVNVLNL 240
Db 181 lpldllnnavatstleelrlyleperwqldledlyrptwqlgkaivfgrksrvvdl 240
QY 241 LKEEVKLYSCTPRNFSVSIREEKRTDTIFWPGLLVKRCGGNACCLHNCNECQCVPK 300
Db 241 lteevrlyscprnfsvsiireelkrtdtifwpgcllvkrccgncacclhncnecqcvpk 300
QY 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
Db 301 vtkeyhevlqlrpktrvgvghnksltdvalehheecdcvcrgstgg 345

RESULT 9
AAB44322

ID AAB44322 standard; Protein; 345 AA.
XX
AC AAB44322;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW expressed sequence tag; detection; cancer.
XX
OS Homo sapiens.
XX
PN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PF 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28585.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavan IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart RA, Tumas D, Williams PM, Wood WI;
XX
XX WPI; 2000-611443/58.
DR N-PSDB; AAC78582.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX
XX Claim 12; Fig 207; 636pp; English.
PS
XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.2e-163;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLTSLALAGORTGTRAESNLSSKQLQSSDKQNGVQDPRHVRVTISGNGSIHS 60
 DB 1 mslfglilltsalagrggtgaesnlsskqfssnkeqngvqdpqheriitvstngsihs 60
 QY 61 PRFPHTYPRNMLVRLVAVDENVRQLTDFDERFGLDEPDICKYDFVEEPPSDGSVL 120
 DB 61 prfphtyprntvrlvaveenvwqltderfgleddedickdfveveepsdgtl 120
 QY 121 GRWCGSTVPKQTSKGNHRIIRFVSDYFPSEPGFCIHYSIIMPOVTEITSPSLVLPSS 180
 DB 121 grwcsgstvpgkqtskgnqirirfvsdeyfpsepgfcihynimpqfteavspvlppsa 180
 QY 181 LSLDLNNAVTAFSTLEELIRYLEPDRQVLDLSLTDALEHHEECDCVCRNAGG 240
 DB 181 lpsldllnnaitafstledliryleperwqldledlyrptwqlgkafgrksrvvdnl 240
 QY 241 LKEEVLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
 DB 241 lteevlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnecqcvpsk 300
 QY 301 VTKKYHEVLQRPKTVGKGLHKSITDVALEHHEECDCVCRNAGG 345
 DB 301 vtckkyhevlqrpktvgkghksitdvalehheecdcvcrgstg 345

RESULT 10

AAB10633

ID AAB10633 standard; Protein; 345 AA.

XX AC

XX AAB10633;

DT 19-JAN-2001 (first entry)

DE Human RACE generated VEGF-X protein.

XX VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.

OS Homo sapiens.

PN WO200037641-A2.

XX PD 29-JUN-2000.

XX PF 21-DEC-1999; 99WO-US30503.

XX PR 22-DEC-1998; 98GB-0028377.

XX PR 18-MAR-1999; 99US-0124967.

XX PR 08-NOV-1999; 99US-0164131.

XX PA (JANC) JANSEN PHARM NV.

XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gostowska A;

XX PI Dhanaraj SN, Xu J;

XX DR WPI; 2000-442669/38.

XX DR N-PSDB; AAA71951.

XX PT New vascular endothelial growth factor protein, useful for treating or

XX PT preventing diseases associated with inappropriate angiogenesis activity
 XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 XX PS Disclosure; Fig 6; 127pp; English.

XX CC This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization

CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the RACE generated human VEGF-X
 CC protein described in the method of the invention.

SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;

Best Local Similarity 87.0%; Pred. No. 1.2e-163;

Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLTSLALAGORTGTRAESNLSSKQLQSSDKQNGVQDPRHVRVTISGNGSIHS 60

DB 1 mslfglilltsalagrggtgaesnlsskqfssnkeqngvqdpqheriitvstngsihs 60

QY 61 PRFPHTYPRNMLVRLVAVDENVRQLTDFDERFGLDEPDICKYDFVEEPPSDGSVL 120

DB 61 prfphtyprntvrlvaveenvwqltderfgleddedickdfveveepsdgtl 120

QY 121 GRWCGSTVPKQTSKGNHRIIRFVSDYFPSEPGFCIHYSIIMPOVTEITSPSLVLPSS 180

DB 121 grwcsgstvpgkqtskgnqirirfvsdeyfpsepgfcihynimpqfteavspvlppsa 180

QY 181 LSLDLNNAVTAFSTLEELIRYLEPDRQVLDLSLTDALEHHEECDCVCRNAGG 240

DB 181 lpsldllnnaitafstledliryleperwqldledlyrptwqlgkafgrksrvvdnl 240

QY 241 LKEEVLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

DB 241 lteevlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnecqcvpsk 300

QY 301 VTKKYHEVLQRPKTVGKGLHKSITDVALEHHEECDCVCRNAGG 345

DB 301 vtckkyhevlqrpktvgkghksitdvalehheecdcvcrgstg 345

RESULT 11

AAB10635

ID AAB10635 standard; Protein; 345 AA.

XX AC

XX AAB10635;

XX DT 19-JAN-2001 (first entry)

XX DE Human VEGF-X protein #1 isolated from clones 4 and 7.

XX VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.

OS Homo sapiens.

XX PN WO200037641-A2.

XX PD 29-JUN-2000.

XX PF 21-DEC-1999; 99WO-US30503.

XX PR 22-DEC-1998; 98GB-0028377.

XX PR 18-MAR-1999; 99US-0124967.

XX PR 08-NOV-1999; 99US-0164131.

```
XX PA (JANC ) JANSSEN PHARM NV.
XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX DR WPI: 2000-442669/38.
XX DR N-PSDB; AAA71955.
XX PT New vascular endothelial growth factor protein, useful for treating or
XX PT preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS Disclosure; Fig 9; 127pp; English.
XX CC This invention describes a novel vascular endothelial growth factor-X
XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX CC antidiabetic activity and acts as an angiogenesis and vascularization
XX CC regulator. An antisense molecule of the invention is useful for treating
XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX CC retinopathy by inhibiting angiogenic activity or inappropriate
XX CC vascularization including formation and proliferation of new blood
XX CC vessels, growth and development of tissues, tissue regeneration and organ
XX CC and tissue repair in a subject. The products of the invention are useful
XX CC for preparing medicaments for treating wounds such as dermal ulcers,
XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote
XX CC skin graft growth, tissue repair, proliferation of new blood vessels,
XX CC tissue regeneration and organ repair by promoting angiogenic activity or
XX CC vascularization. This sequence represents the human VEGF-X protein
XX CC isolated from clones 4 and 7 described in the method of the invention.
XX SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.2e-163;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

Qy 1 MLLGLLLLSALAGQRTGTRAESNLSSKQLQSSDKQNGVQDPRHERVVTISGNGSIHS 60
Db 1 mslfgllltsalagrggtqaesnlsskfssnkeqngvqdpqheriitvstngsihs 60

Qy 61 PKPHTYPRNMVLYWRLVAVDENVRIOITFDERFGLDEPDIDCKYDFVEVEEPSDGSVL 120
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
61 prfphtyprntvlwrlvaveenvvqltfderrfgledpedddickdyfveveepsdgtl 120

Qy 121 GRWCGSGTVPKQTSKGNHIRFVSDSEYFPSEPGFCIHYSIMPVQVTTTSPSVLPSS 180
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
121 grwcsgstvpqkqskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180

Qy 181 LSLDLLNNAVTAFTLEELIRYLEPDRQVLDLSLYKPTWOLLGKAFLYGKSKVYNLNL 240
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
181 lpsldllnnatfstledliryleperwqldeledyrtwqlgkafvgrksrvdlnl 240

RESULT 12
AAB10644
ID AAB10644 standard; Protein; 345 AA.
XX AC AAB10644;
XX DT 19-JAN-2001 (first entry)
XX DE Human VEGF-X protein #4.
```

```
XX KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
XX KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
XX KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
XX KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
XX KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
XX KW venous sore; diabetic ulcer; burns; skin graft growth.
XX OS Homo sapiens.
XX PN WO200037641-A2.
XX XX 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-US30503.
XX PR 22-DEC-1998; 98GB-0028377.
XX PR 18-MAR-1999; 99US-0124967.
XX PR 08-NOV-1999; 99US-0164131.
XX (JANC ) JANSSEN PHARM NV.
XX PA Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX DR WPI: 2000-442669/38.
XX DR N-PSDB; AAA71990.
XX PT New vascular endothelial growth factor protein, useful for treating or
XX PT preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS Disclosure; Fig 30B; 127pp; English.
XX CC This invention describes a novel vascular endothelial growth factor-X
XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX CC antidiabetic activity and acts as an angiogenesis and vascularization
XX CC regulator. An antisense molecule of the invention is useful for treating
XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX CC retinopathy by inhibiting angiogenic activity or inappropriate
XX CC vascularization including formation and proliferation of new blood
XX CC vessels, growth and development of tissues, tissue regeneration and organ
XX CC and tissue repair in a subject. The products of the invention are useful
XX CC for preparing medicaments for treating wounds such as dermal ulcers,
XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote
XX CC skin graft growth, tissue repair, proliferation of new blood vessels,
XX CC tissue regeneration and organ repair by promoting angiogenic activity or
XX CC vascularization. This sequence represents a human VEGF-X protein
XX CC described in the method of the invention.
XX SQ Sequence 345 AA;

Query Match 90.2%; Score 1667; DB 21; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.2e-163;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

Qy 1 MLLGLLLLSALAGQRTGTRAESNLSSKQLQSSDKQNGVQDPRHERVVTISGNGSIHS 60
Db 1 mslfgllltsalagrggtqaesnlsskfssnkeqngvqdpqheriitvstngsihs 60

Qy 61 PKPHTYPRNMVLYWRLVAVDENVRIOITFDERFGLDEPDIDCKYDFVEVEEPSDGSVL 120
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
61 prfphtyprntvlwrlvaveenvvqltfderrfgledpedddickdyfveveepsdgtl 120

Qy 121 GRWCGSGTVPKQTSKGNHIRFVSDSEYFPSEPGFCIHYSIMPVQVTTTSPSVLPSS 180
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
121 grwcsgstvpqkqskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180

Qy 181 LSLDLLNNAVTAFTLEELIRYLEPDRQVLDLSLYKPTWOLLGKAFLYGKSKVYNLNL 240
Db 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
181 lpsldllnnatfstledliryleperwqldeledyrtwqlgkafvgrksrvdlnl 240
```


Qy 1 MLLGLLLLTSLAQRTGTRAESNLSSKLQSSDKEQNGVODPRHERVVTISGNGSIHS 60
Db 1 mslfgllltsalagrgtgaesnlsskfgfssnkegvgdpgqgherlitvstngsihs 60
Qy 61 PKPHTYPRNMVLRVAVDENVRIOLEDFRFGLEDPEDDICKYDFVEVEEPSDGSVL 120
Db 61 prfptyprntvwrivaveenvwiqitfderfgledpeddickdyfveveepsdgtll 120
Qy 121 GRWCGSTVPCKQTSKGNHIRIRFVSDYFSEPGFCIHYSIIMPQVTTETSPSVLPSS 180
Db 121 grwsgstvpgkqiskgnqirirfvsdeyfpsepfcihynivmpqfteavspsvlpssa 180
Qy 181 LSLDLLNNAVTAFTLEELIRYLEPDRQVLDLSYKPTWQLLGRAFLYGGKSKVVNLNL 240
Db 181 lpdllnnaitafstledliryleperwqldledlyrptwqlkgafvgrksrvvdlnl 240
Qy 241 LKEEVKLYSCTPRNFSVSIRELKRDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 lteevrlyscprnfsvsireelkrtdtlfwpgcllvkrcgncacclhncnecqcvpsk 300
Qy 301 VTKKYHEVLQLRPKTGKGLHKSITDVALEHHEECDCVCRGNAGG 345
Db 301 vtkeyhevlqlrpkgtgvrghksltdvalehheecdcvcrgstgg 345

Search completed: September 5, 2001, 10:54:33
Job time: 30 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein : protein search, using sw model

Run on: September 5, 2001, 10:55:18 ; Search time 16 Seconds
(without alignments)
1642.514 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 345
Sequence: 1 MLLGLLLLTALAGQRTGT.....DVALEHHECDVCGRNAGG 345

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.6	97	4 S26485	cytochrome P450 21
2	9	2.6	494	1 O4HUC2	steroid 21-monooxy
3	8	2.3	33	2 A36154	benzphetamine N-de
4	8	2.3	105	2 A26774	platelet factor 4
5	8	2.3	124	2 I34768	epididymis-specifi
6	8	2.3	125	2 S04936	Ig kappa chain pre
7	8	2.3	130	2 P10113	Ig kappa chain pre
8	8	2.3	221	2 D82294	DNA mismatch repai
9	8	2.3	269	2 T04095	ribonuclease S hom
10	8	2.3	373	2 S43455	hypothetical prote
11	8	2.3	380	2 A83458	hypothetical prote
12	8	2.3	461	2 A82220	hypothetical prote
13	8	2.3	496	2 F75257	hypothetical prote
14	8	2.3	526	2 G86587	heat shock protein
15	8	2.3	526	2 D72036	heat shock protein
16	8	2.3	526	2 F81504	60 kDa chaperonin
17	8	2.3	662	2 D40228	neurexin II-beta p
18	8	2.3	707	2 JC2218	procollagen C-endo
19	8	2.3	730	1 BMHUI	procollagen C-endo
20	8	2.3	823	1 A58788	procollagen C-endo
21	8	2.3	874	2 J00881	genome polypeptid
22	8	2.3	986	1 B58788	procollagen C-endo
23	8	2.3	991	2 I49540	procollagen C-endo
24	8	2.3	994	2 I49276	c-mer tyrosine kin
25	8	2.3	999	1 I3HUG3	desmoglein 3 precu
26	8	2.3	3033	1 J01303	genome polypeptid
27	7	2.0	68	2 I49136	dopamine transport
28	7	2.0	85	2 A53453	gonadoliberin II p
29	7	2.0	90	1 D69782	conserved hypothet

30 7 2.0 100 2 G83587 hypothetical prote
31 7 2.0 116 2 A27594 Ig kappa chain pre
32 7 2.0 117 1 K1HU11 Ig kappa chain pre
33 7 2.0 117 1 K1HU12 Ig kappa chain pre
34 7 2.0 117 2 C21056 Ig kappa chain pre
35 7 2.0 117 2 B21056 Ig kappa chain pre
36 7 2.0 117 2 S10237 Ig kappa chain pre
37 7 2.0 117 2 S42263 Ig kappa chain V r
38 7 2.0 117 2 S11700 Ig kappa chain pre
39 7 2.0 117 2 S24206 Ig kappa chain V r
40 7 2.0 117 2 S24207 Ig kappa chain V r
41 7 2.0 117 2 S41809 Ig kappa chain V r
42 7 2.0 117 2 S41810 Ig kappa chain V r
43 7 2.0 117 2 S41811 Ig kappa chain V r
44 7 2.0 117 2 S41812 Ig kappa chain V r
45 7 2.0 117 2 S42264 Ig kappa chain V r

ALIGNMENTS

RESULT 1

S26485
cytochrome P450 21A/B mutant fusion protein - human
N;Alternate names: steroid 21-monooxygenase
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 17-Aug-1995 #text_change 15-Feb-1996
C;Accession: S26485; S29672
R;Helmberg, A.; Kofler, R.
submitted to the EMBL Data Library, March 1991
A;Reference number: S26484
A;Accession: S26485
A;Molecule type: DNA
A;Residues: 1-97 <HEL>
A;Cross-references: EMBL:X58901
A;Experimental source: leukocyte clone AGS 8-23
A;Note: an unequal cross-over mutation of the CYP21P pseudogene and CYP21 gene in a c
A;Accession: S29672
A;Molecule type: DNA
A;Residues: 1-97 <HE2>
A;Cross-references: EMBL:X58908
C;Genetics:
A;Gene: CYP21P/CYP21
A;Map position: 6p21.3
A;Introns: 68/1
C;Keywords: fusion protein

Query Match 2.6%; Score 9; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLGLLLLL 9

|||||

Db 1 MLLGLLLLL 9

RESULT 2

O4HUC2

steroid 21-monooxygenase (EC 1.14.99.10) cytochrome P450 21 - human
N;Alternate names: cytochrome P450(C21B); steroid 21-hydroxylase cytochrome P450 21A2
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 08-Feb-1996 #text_change 03-Mar-2000
C;Accession: A25446; A00191; A27865; A32715; A21889; S26484; S29670; S29671; S26584;
R;White, P.C.; New, M.I.; Dupont, B.
Proc. Natl. Acad. Sci. U.S.A. 83, 5111-5115, 1986
A;Title: Structure of human steroid 21-hydroxylase genes.
A;Reference number: A94108; MUID:86259742
A;Accession: A25446
A;Molecule type: DNA

A;Residues: 1-494 <WHI>
A;Cross-references: GB:M13936; NID:g187899; PIDN:AAA59695.1; PID:g386910
R;Higashi, Y.; Yoshioka, H.; Yamane, M.; Gotoh, O.; Fujii-Kuriyama, Y.

Proc. Natl. Acad. Sci. U.S.A. 83, 2841-2845, 1986
A:Title: Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged
A:Reference number: A00191; MUID:86206051
A:Accession: A00191
A:Molecule type: mRNA
A:Residues: 1-425, 'P', 427-494 <HIG>
R:Rodrigues, N.R.; Dunham, I.; Yu, C.Y.; Carroll, M.C.; Porter, R.R.; Campbell, R.D.
EMBO J. 6, 1653-1661, 1987
A:Title: Molecular characterization of the HLA-linked steroid 21-hydroxylase B gene from
A:Reference number: A27865; MUID:87275858
A:Accession: A27865
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-9, 'L', 10-101, 'R', 103-372, 'S', 373-494 <ROD>
R:Matteson, K.J.; Phillips III, J.A.; Miller, W.L.; Chung, B.; Orlando, P.J.; Frisch, H.
Proc. Natl. Acad. Sci. U.S.A. 84, 5858-5862, 1987
A:Title: P450XXI (steroid 21-hydroxylase) gene deletions are not found in family studies
A:Reference number: A32715; MUID:87289701
A:Accession: A32715
A:Molecule type: mRNA
A:Residues: 265-310, 'L', 312-345, 'I', 347-494 <MAT>
A:Cross-references: GB:M1725; NID:g189446; PIDN:AAA59985.1; PID:g386993
A:Note: the authors translated the codon ATT for residue 346 as Asn
R:Carroll, M.C.; Campbell, R.D.; Porter, R.R.
Proc. Natl. Acad. Sci. U.S.A. 82, 521-525, 1985
A:Title: Mapping of steroid 21-hydroxylase genes adjacent to complement component C4 gene
A:Reference number: A21889; MUID:85113228
A:Accession: A21889
A:Molecule type: DNA
A:Residues: 149-171, 'N', 173-182 <CAR>
A:Cross-references: GB:K02771; NID:g187928; PIDN:AAA59706.1; PID:g443672
R:Helmsberg, A.; Kofler, R.
submitted to the EMBL Data Library, March 1991
A:Reference number: S26484
A:Accession: S26484
A:Molecule type: DNA
A:Residues: 1-101, 'R', 103-371 <HEL>
A:Cross-references: EMBL:X58904; NID:g30319; PIDN:CAA41707.1; PID:g30320
A:Accession: S29670
A:Molecule type: DNA
A:Residues: 1-9, 'L', 10-101, 'R', 103-371 <HE4>
A:Cross-references: EMBL:X58902; NID:g30325; PIDN:CAA41706.1; PID:g30326
R:Helmsberg, A.; Tabarelli, M.; Dobler, G.; Kofler, R.
submitted to the EMBL Data Library, March 1991
A:Description: Identification of molecular defects causing congenital adrenal hyperplasia
A:Reference number: S26584
A:Accession: S26584
A:Molecule type: DNA
A:Residues: 1-171, 'N', 173-371 <HE2>
A:Cross-references: EMBL:X58988; NID:g30316; PIDN:CAA41702.1; PID:g30317
A:Accession: S29673
A:Molecule type: DNA
A:Residues: 1-9, 'L', 10-101, 'R', 103-371 <HE5>
A:Cross-references: EMBL:X58900; NID:g30328; PIDN:CAA41704.1; PID:g30329
R:Globerman, H.; Amor, M.; Parker, K.L.; New, M.I.; White, P.C.
J. Clin. Invest. 82, 139-144, 1988
A:Title: Nonsense mutation causing steroid 21-hydroxylase deficiency.
A:Reference number: 155547; MUID:88273565
A:Accession: 155547
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9, 'L', 10-280, 'L', 282-494 <RES>
A:Cross-references: GB:M21550; NID:g180960; PIDN:AAA52063.1; PID:g180962
R:Amor, M.; Parker, K.L.; Globerman, H.; New, M.I.; White, P.C.
Proc. Natl. Acad. Sci. U.S.A. 85, 1600-1604, 1988
A:Title: Mutation in the CYP21B gene (ile-172----Asn) causes steroid 21-hydroxylase deficiency
A:Reference number: I59109; MUID:88144483
A:Accession: I59109
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 149-182 <RE2>
A:Cross-references: GB:M19711; NID:g181289; PIDN:AAA83248.1; PID:g181290
R:Collier, S.; Tassabehji, M.; Strachan, T.
Nature Genet. 3, 260-265, 1993
A:Title: A de novo pathological point mutation at the 21-hydroxylase locus: implicati
A:Reference number: I58113; MUID:93251047
A:Accession: I58113
A:Molecule type: DNA
A:Residues: 109-171, 'N', 173-185 <RE3>
A:Cross-references: GB:S60612; NID:g300314
C:Comment: Deficiency of this enzyme (21-hydroxylase deficiency) causes about 90 % of
C:Genetics:
A:Gene: CYP21; CYP21B
A:Cross-references: GDB:120605; OMIM:201910
A:Map position: 6p21.3-6p21.3
A:Introns: 67/1; 97/1; 148/3; 182/3; 216/3; 245/3; 312/3; 372/2; 407/1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F:288-450/Domain: cytochrome P450 homology <CYP>
F:428/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 2.6%; Score 9; DB 1; Length 494;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLGLLLLL 9
| | | | | | | | | |
DB 1 MLLGLLLLL 9
RESULT 3
A36154
benzphetamine N-demethylase (EC 1.14.14.-) cytochrome P450 2B - guinea pig (fragment)
N:Alternate names: cytochrome P450 (GP-1)
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Mar-1992 #sequence revision 08-Feb-1996 #text_change 05-Mar-1999
C:Accession: S15135; S28205; A36154
R:Oguri, K.; Kaneko, H.; Tanimoto, Y.; Yamada, H.; Yoshimura, H.
Arch. Biochem. Biophys. 287, 105-111, 1991
A:Title: A constitutive form of guinea pig liver cytochrome P450 closely related to p
A:Reference number: S15135; MUID:91378369
A:Accession: S15135
A:Molecule type: protein
A:Residues: 1-33 <ARC>
R:Yamada, H.; Kaneko, H.; Takeuchi, K.; Oguri, K.; Yoshimura, H.
Arch. Biochem. Biophys. 299, 248-254, 1992
A:Title: Tissue-specific expression, induction, and inhibition through metabolic inte
A:Reference number: S28205; MUID:93073973
A:Accession: S28205
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-31 <YAM>
R:Narimatsu, S.; Akutsu, Y.; Matsunaga, T.; Watanabe, K.; Yamamoto, I.; Yoshimura, H.
Biochem. Biophys. Res. Commun. 172, 607-613, 1990
A:Title: Purification of a cytochrome P450 isozyme belonging to a subfamily of P450 I
A:Reference number: A36154; MUID:91054472
A:Accession: A36154
A:Molecule type: protein
A:Residues: 1-20 <NAR>
C:Genetics:
A:Gene: CYP2B
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; endoplasmic reticulum; heme; monooxygenase; oxidoreduc
Query Match 2.3%; Score 8; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LLLGLLLLL 9

Db 11 LLLGLLLL 18
|||||

RESULT 4

A26774
platelet factor 4 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Aug-1999
C:Accession: A26774; S45657
A:Doi, T.; Greenberg, S.M.; Rosenberg, R.D.
Mol. Cell. Biol. 7, 898-904, 1987
A:Title: Structure of the rat platelet factor 4 gene: a marker for megakaryocyte differentiation
A:Reference number: A26774; MUID:87144262
A:Accession: A26774
A:Molecule type: DNA; mRNA
A:Residues: 1-105 <DOI>
A:Cross-references: GB:M15254; NID:g206090; PIDN:AAA41832.1; PID:g206091
R:Ravanat, C.; Gachet, C.; Herbert, J.M.; Schuhler, S.; Guilleminot, J.C.; Uzabiaga, F.; F
Eur. J. Biochem. 223, 203-210, 1994
A:Title: Rat platelets contain glycosylated and non-glycosylated forms of platelet factor 4
A:Reference number: S45657; MUID:94307262
A:Accession: S45657
A:Molecule type: protein
A:Residues: 30-42 <RAV>
C:Superfamily: beta-thromboglobulin

Query Match 2.3%; Score 8; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
|||||

Db 17 LLLGLLLL 24
|||||

RESULT 5

I54768
epididymis-specific four-disulfide core protein CE4 - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 24-Oct-2000
C:Accession: I54768
R:Ellerbrock, K.; Pera, I.; Hartung, S.; Ivell, R.
Int. J. Androl. 17, 314-323, 1994
A:Title: Gene expression in the dog epididymis: a model for human epididymal function.
A:Reference number: I54768; MUID:95263175
A:Accession: I54768
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-124 <ELL>
A:Cross-references: GB:S77395; NID:g945180; PIDN:AAB34264.1; PID:g945181
C:Superfamily: antileukoproteinase; antileukoproteinase repeat homology
F:76-123/Domain: antileukoproteinase repeat homology <ALP>

Query Match 2.3%; Score 8; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
|||||

Db 13 LLLGLLLL 20
|||||

RESULT 6

Ig kappa chain precursor V-J region (IHL) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
C:Accession: S04936
R:Levy, S.; Mendel, E.; Kon, S.; Avnur, Z.; Levy, R.
J. Exp. Med. 168, 475-489, 1988

A:Title: Mutational hot spots in Ig V region genes of human follicular lymphomas.
A:Reference number: S04936; MUID:88316166
A:Accession: S04936
A:Molecule type: mRNA
A:Residues: 1-125 <LEV>
A:Cross-references: EMBL:X13076; NID:g33173; PIDN:CAA31477.1; PID:g736243
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-17/Domain: signal sequence (fragment) #status predicted <SIG>
F:18-125/Product: Ig kappa chain (fragment) #status predicted <MAT>
F:33-107/Domain: immunoglobulin homology <IMM>

Query Match 2.3%; Score 8; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
|||||

Db 3 LLLGLLLL 10
|||||

RESULT 7

PL0113
Ig kappa chain precursor V-I region (CJ) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jan-2000
C:Accession: PL0113
R:Levy, S.; Mendel, E.; Kon, S.; Avnur, Z.; Levy, R.
J. Exp. Med. 168, 475-489, 1988
A:Title: Mutational hot spots in Ig V region genes of human follicular lymphomas.
A:Reference number: S04936; MUID:88316166
A:Accession: PL0113
A:Molecule type: mRNA
A:Residues: 1-130 <LEV>
A:Experimental source: follicular lymphoma cells
A:Note: the sequence shown here is derived from the consensus nucleotide sequence of
om tumor cells of a single patient
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-130/Product: Ig kappa chain V-I region CJ #status predicted <MAT>
F:38-112/Domain: immunoglobulin homology <IMM>
F:46-56/Region: complementarity-determining 1
F:72-78/Region: complementarity-determining 2
F:111-130/Region: complementarity-determining 3
F:118-130/Region: J1

Query Match 2.3%; Score 8; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
|||||

Db 8 LLLGLLLL 15
|||||

RESULT 8

D82294
DNA mismatch repair protein Mth VC0668 [imported] - Vibrio cholerae (strain N16961 s
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82294
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833
A:Accession: D82294
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-221 <HEI>

A:Cross-references: GB:AE004153; GB:AE003852; NID:g9655103; PIDN:AAF93833.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0668
A:Map position: 1
C:Superfamily: mutator muth

Query Match 2.3%; Score 8; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 307 EVLQLRPK 314
|||||||
Db 173 EVLQLRPK 180

RESULT 9

T04095
ribonuclease S homolog - maize
N:Alternate names: S-like RNase
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C:Accession: T04095
R:Char, B.R.; Hake, S.
submitted to the EMBL Data Library, August 1996
A:Reference number: 215206
A:Accession: T04095
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-269 <CHA>
A:Cross-references: EMBL:U66241; NID:g1698669; PIDN:AAB37265.1; PID:g1698670
C:Genetics:
A:Gene: knl
A:Map position: 7L

Query Match 2.3%; Score 8; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLL 9
|||||||
Db 8 LLLGLLLL 15

RESULT 10

S43455
hypothetical protein (LEU2 3' region) - yeast (Pichia angusta) (fragment)
C:Species: Pichia angusta
C:Date: 13-Jan-1995 #sequence_revision 23-Feb-1996 #text_change 20-Apr-2000
C:Accession: S43455
R:Agaphonov, M.O.; Poznyakovski, A.I.; Bogdanova, A.I.; Ter-Avanesyan, M.D.
yeast 10, 509-513, 1994
A:Title: Isolation and characterization of the LEU2 gene of Hansenula polymorpha.
A:Reference number: S43454; MUID:95028149
A:Accession: S43455
A:Molecule type: DNA
A:Residues: 1-373 <AGA>
A:Cross-references: EMBL:U00889; NID:g392892; PIDN:AAA19110.1; PID:g392894
C:Genetics:
A:Map position: 1

Query Match 2.3%; Score 8; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLL 9
|||||||
Db 270 LLLGLLLL 277

RESULT 11

A83458
hypothetical protein PA1509 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83458
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lozy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: A83458
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <STO>
A:Cross-references: GB:AE004579; GB:AE004091; NID:g9947455; PIDN:AAG04898.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1509

Query Match 2.3%; Score 8; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 PSEPGFCI 158
|||||||
Db 174 PSEPGFCI 181

RESULT 12

A82220
hypothetical protein VC1265 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82220
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: A82220
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <HEI>
A:Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94424.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1265
A:Map position: 1

Query Match 2.3%; Score 8; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLLLTSAI 13
|||||||
Db 7 LLLLTSAI 14

RESULT 13

F75257
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75257
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: F75257
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-496 <WHI>
 A:Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF12116.1; PID:g6460404
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2572
 A:Map position: 1

Query Match 2.3%; Score 8; DB 2; Length 496;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LLLLSAL 13

|||||||

Db 304 LLLLSAL 311

RESULT 14

G86587

heat shock protein-60 [imported] - Chlamydomophila pneumoniae (strain J138)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: G86587

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349

A:Accession: G86587

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-526 <STO>

A:Cross-references: GB:BA000008; NID:g8979150; PIDN:BA09985.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: groEL2

C:Superfamily: chaperonin groEL

Query Match

2.3%; Score 8; DB 2; Length 526;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 EELIRYLE 204

|||||||

Db 228 EELIRYLE 235

RESULT 15

D72036

heat shock protein-60 - Chlamydomophila pneumoniae (strain CWL029)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000

C:Accession: D72036

R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Pan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606

A:Accession: D72036

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-526 <ARN>

A:Cross-references: GB:AE001659; GB:AE001363; NID:g4377081; PIDN:AAD18915.1; PID:g437708

A:Experimental source: strain CWL029

C:Genetics:

A:Gene: groEL2

C:Superfamily: chaperonin groEL

Query Match 2.3%; Score 8; DB 2; Length 526;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 EELIRYLE 204

|||||||

Db 228 EELIRYLE 235

Search completed: September 5, 2001, 10:57:14

Job time: 116 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:03 ; Search time 12.27 Seconds
(without alignments)
578.946 Million cell updates/sec

Title: us-09-457-066-43
Perfect score: 1848
Sequence: 1 MLLGLLLLTALAGORTGT.....DVALEHHECDVCVRNAGG 345

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	181	9.8	788	1	US-08-572-225-1
2	169	9.1	923	2	US-08-936-135-6
3	165	8.9	1013	2	US-08-866-650-3
4	165	8.9	1013	2	US-09-021-287-3
5	162	8.8	591	3	US-08-991-408-4
6	162	8.8	1013	2	US-08-866-650-5
7	162	8.8	1013	2	US-09-021-287-5
8	162	8.8	1013	3	US-08-991-408-2
9	158	8.5	449	2	US-08-839-008-2
10	158	8.5	449	2	US-08-839-008-9
11	155.5	8.4	901	3	US-08-936-135-22
12	155.5	8.4	906	3	US-08-936-135-24
13	155.5	8.4	909	3	US-08-936-135-8
14	155.5	8.4	909	3	US-08-936-135-10
15	155.5	8.4	909	3	US-08-936-135-18
16	155.5	8.4	914	3	US-08-936-135-12
17	155.5	8.4	926	3	US-08-936-135-14
18	155.5	8.4	926	3	US-08-936-135-20
19	155.5	8.4	931	3	US-08-936-135-16
20	153.5	8.3	415	4	US-09-032-523-2
21	136	7.4	325	4	US-08-915-795-3
22	136	7.4	354	4	US-08-915-795-5
23	135.5	7.3	358	4	US-08-915-795-8
24	133.5	7.2	401	2	US-08-839-008-5
25	133.5	7.2	468	2	US-08-839-008-7
26	133.5	7.2	468	4	US-09-032-523-8
27	132	7.1	321	4	US-08-915-795-9

Sequence 11, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 18, Appl
Sequence 8, Appl
Sequence 35, Appl
Sequence 2, Appl
Sequence 13, Appl
Sequence 2, Appl
Sequence 27, Appl
Sequence 37, Appl
Sequence 27, Appl

28 132 7.1 415 4 US-08-795-430-11
29 128 6.9 277 1 US-08-024-868-2
30 128 6.9 277 2 US-08-242-097-2
31 128 6.9 277 4 US-09-206-695-2
32 128 6.9 277 5 PCT-US96-11995-1
33 127.5 6.9 101 2 US-08-242-097-6
34 127.5 6.9 101 4 US-09-206-695-6
35 127 6.9 419 2 US-08-999-811-2
36 127 6.9 419 3 US-09-042-105-2
37 127 6.9 419 3 US-09-042-105-18
38 127 6.9 419 4 US-08-795-430-8
39 127 6.9 419 4 US-08-510-133A-35
40 121 6.5 419 5 PCT-US96-09001-2
41 117.5 6.4 418 4 US-08-795-430-13
42 117.5 6.4 1290 1 US-08-470-350B-2
43 115.5 6.2 205 3 US-08-989-251-27
44 115.5 6.2 205 3 US-08-989-251-37
45 115.5 6.2 205 3 US-09-340-250-27

ALIGNMENTS

RESULT 1
US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,225
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-572-225-1

Query Match 9.8%; Score 181; DB 1; Length 788;
Best Local Similarity 42.6%; Pred. No. 6.7e-11;
Matches 46; Conservative 16; Mismatches 38; Indels 8; Gaps 5;

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: US/09/021,287
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/866,650
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Berson, Bennett J
;; REGISTRATION NUMBER: 37094
;; REFERENCE/DOCKET NUMBER: 960296.93839
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-251-5000
;; TELEFAX: 608-251-9166
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1013 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-021-287-3

Query Match 8.9%; Score 165; DB 2; Length 1013;
Best Local Similarity 40.7%; Pred. No. 5.8e-09;
Matches 44; Conservative 21; Mismatches 35; Indels 8; Gaps 7;

QY 55 NGSIHSPKFPHTYPRNMVLRVAVDENVRIQLTFDERFGLDEPDICKYDFVEV-EE 113
DB 626 NGTITPGWPKFPPKNCVQVIAFSQ-YRISKV-FELEG-NEVCKYDVEIWSG 681
QY 114 PSDGSVLGRWCGSTVPKQTSKGNHIRFVSDYFPPSEPGFCIH 160
DB 682 PSSEKLGKFCGA-DIPEVMTSHFNNMRIFKSDNTV-SKKGFKAHF 727

RESULT 5
US-08-991-408-4
; Sequence 4, Application US/08991408
; Patent No. 6008017
; GENERAL INFORMATION:
; APPLICANT: ARLETH, ANTHONY J.
; APPLICANT: WILLETTTE, ROBERT N.
; APPLICANT: ELSHOURBAGY, NABIL A.
; APPLICANT: LI, XIAOTONG
; TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,408
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/034,471
; FILING DATE: 02-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: ATG-50038
; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 610-407-0700
;; TELEFAX: 610-407-0701
;; TELEX: 846169
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 591 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-991-408-4

Query Match 8.8%; Score 162; DB 3; Length 591;
Best Local Similarity 39.8%; Pred. No. 5.2e-09;
Matches 43; Conservative 22; Mismatches 35; Indels 8; Gaps 6;

QY 55 NGSIHSPKFPHTYPRNMVLRVAVDENVRIQLTFDERFGLDEPDICKYDFVEE- 113
DB 204 NGTITPGWPKFPPKNCVQVIAFSQ-YRISKV-FELEG-NEVCKYDVEIWSG 259
QY 114 -PSDGSVLGRWCGSTVPKQTSKGNHIRFVSDYFPPSEPGFCIH 160
DB 260 LSSEKLGKFCGA-EVPEVITSOFNMRIFKSDNTV-SKKGFKAHF 305

RESULT 6
US-08-866-650-5
; Sequence 5, Application US/08866650
; Patent No. 5939321
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-866-650-5

Query Match 8.8%; Score 162; DB 2; Length 1013;
Best Local Similarity 39.8%; Pred. No. 1.2e-08;
Matches 43; Conservative 22; Mismatches 35; Indels 8; Gaps 6;
QY 55 NGSIHSPKFPHTYPRNMVLRVAVDENVRIQLTFDERFGLDEPDICKYDFVEE- 113

```
Db 626 NGTITPGWPKPEYPPNKNKVQVWVAPTQ-YRISVKF-EFFELEG--NEVCKYDYVEIWSG 681
Qy 114 -PSDGSVLGRWCGSGTVPGKQTSKGNHIRIRFVSDYFPPSEPGFCIH 160
Db 682 LSSESKLHGKFCGA-EVPEVITSQFNMMRIEFSKSDNTV-SKKGFKAHF 727

RESULT 7
US-09-021-287-5
; Sequence 5, Application US/09021287
; Patent No. 5981717
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/866,650
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-021-287-5

Query Match 8.8%; Score 162; DB 2; Length 1013;
Best Local Similarity 39.8%; Pred. No. 1.2e-08;
Matches 43; Conservative 22; Mismatches 35; Indels 8; Gaps 6;

Qy 55 NGSIHSPKFPHTYPRNMVLVRLVAVDENVRIQLTFDERFGLDEPDICKYDFVEVEE- 113
Db 626 NGTITPGWPKPEYPPNKNKVQVWVAPTQ-YRISVKF-EFFELEG--NEVCKYDYVEIWSG 681
Qy 114 -PSDGSVLGRWCGSGTVPGKQTSKGNHIRIRFVSDYFPPSEPGFCIH 160
Db 682 LSSESKLHGKFCGA-EVPEVITSQFNMMRIEFSKSDNTV-SKKGFKAHF 727

RESULT 8
US-08-991-408-2
; Sequence 2, Application US/08991408
; Patent No. 6008017
; GENERAL INFORMATION:
; APPLICANT: ARLETH, ANTHONY J.
; APPLICANT: WILLETTTE, ROBERT N.
```

```
; APPLICANT: ELSHOUBAGY, NABIL A.
; APPLICANT: LI, XIAOTONG
; TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,408
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/034,471
; FILING DATE: 02-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: ATG-50038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-991-408-2

Query Match 8.8%; Score 162; DB 3; Length 1013;
Best Local Similarity 39.8%; Pred. No. 1.2e-08;
Matches 43; Conservative 22; Mismatches 35; Indels 8; Gaps 6;

Qy 55 NGSIHSPKFPHTYPRNMVLVRLVAVDENVRIQLTFDERFGLDEPDICKYDFVEVEE- 113
Db 626 NGTITPGWPKPEYPPNKNKVQVWVAPTQ-YRISVKF-EFFELEG--NEVCKYDYVEIWSG 681
Qy 114 -PSDGSVLGRWCGSGTVPGKQTSKGNHIRIRFVSDYFPPSEPGFCIH 160
Db 682 LSSESKLHGKFCGA-EVPEVITSQFNMMRIEFSKSDNTV-SKKGFKAHF 727

RESULT 9
US-08-839-008-2
; Sequence 2, Application US/08839008
; Patent No. 5916758
; GENERAL INFORMATION:
; APPLICANT: Hurlle, Mark R
; APPLICANT: McDonnell, Peter C
; APPLICANT: McNulty, Dean E
; APPLICANT: Rosen, Craig A
; APPLICANT: Siemens, Ivo R
; APPLICANT: Young, Peter R
; APPLICANT: Yue, Tian-Li
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
```


TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 901 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-22

Query Match 8.4%; Score 155.5; DB 3; Length 901;
Best Local Similarity 32.8%; Pred. No. 5.3e-08;
Matches 43; Conservative 18; Mismatches 63; Indels 7; Gaps 4;
QY 34 SDKEQGVQDPRHERVVTISGNGSIHSPKFPHTYPRNMVLVRLVAVDENVRIQLTFDER 93
Db 16 SGHEVRSQQDPPCGGRPNKSDAGYITSPGYPDYPSHONCEWIVYAPEPNQKIVLNFNP 75
QY 94 FGLEDPEDDICKYDFVEVEEPSDGS--VLGRWCGSGTVPGKOTSGNHNIRIRFVSDEYFP 151
Db 76 FEIEKHD---CKYDFIEIRGDSEADLLGKHCGN-IAPPTIISGSLVYIKFTSD-YAR 130
QY 152 SEPGFCIHYSI 162
Db 131 QGAGFSLRYEI 141

RESULT 12
US-08-936-135-24
Sequence 24, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-24

Query Match 8.4%; Score 155.5; DB 3; Length 906;
Best Local Similarity 32.8%; Pred. No. 5.4e-08;
Matches 43; Conservative 18; Mismatches 63; Indels 7; Gaps 4;
QY 34 SDKEQGVQDPRHERVVTISGNGSIHSPKFPHTYPRNMVLVRLVAVDENVRIQLTFDER 93
Db 16 SGHEVRSQQDPPCGGRPNKSDAGYITSPGYPDYPSHONCEWIVYAPEPNQKIVLNFNP 75
QY 94 FGLEDPEDDICKYDFVEVEEPSDGS--VLGRWCGSGTVPGKOTSGNHNIRIRFVSDEYFP 151
Db 76 FEIEKHD---CKYDFIEIRGDSEADLLGKHCGN-IAPPTIISGSLVYIKFTSD-YAR 130
QY 152 SEPGFCIHYSI 162
Db 131 QGAGFSLRYEI 141

RESULT 13
US-08-936-135-8
Sequence 8, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 909 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-8

Query Match 8.4%; Score 155.5; DB 3; Length 909;
Best Local Similarity 32.8%; Pred. No. 5.4e-08;
Matches 43; Conservative 18; Mismatches 63; Indels 7; Gaps 4;
QY 34 SDKEQGVQDPRHERVVTISGNGSIHSPKFPHTYPRNMVLVRLVAVDENVRIQLTFDER 93
Db 16 SGHEVRSQQDPPCGGRPNKSDAGYITSPGYPDYPSHONCEWIVYAPEPNQKIVLNFNP 75
QY 94 FGLEDPEDDICKYDFVEVEEPSDGS--VLGRWCGSGTVPGKOTSGNHNIRIRFVSDEYFP 151
Db 76 FEIEKHD---CKYDFIEIRGDSEADLLGKHCGN-IAPPTIISGSLVYIKFTSD-YAR 130
QY 152 SEPGFCIHYSI 162



A58788
procollagen C-endopeptidase (EC 3.4.24.19) precursor, splice form HIS - human
N;Alternate names: bone morphogenic protein splice form BMP-1/His
C;Species: Homo sapiens (man)
C;Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 18-Jun-1999
C;Accession: A37278; A58788

R;Wozney, J.M.; Rosen, V.: Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew Science 242, 1528-1534, 1988

A>Title: Novel regulators of bone formation: molecular clones and activities.

A:Reference number: A37278; PMID:89072730

A:Accession: A37278

A:Molecule type: mRNA

A:Residues: 1-702; EKRPAOPPGRGHQLKPRVOKNRNTPQ' <WOZ>

A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500

J.Takahara, K.; Lyons, G.E.; Greenspan, D.S.

F: Biol. Chem. 269, 32572-32578, 1994

A>Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by two genes from mouse embryo.

A:Reference number: A58788; MOID:95096114

A:Accession: A58788

A:Molecule type: mRNA

A:Residues: 703-823 <TAU>

A:Cross-references: GB:L35278; NID:g619423; PIDN:AAC41703.1; PID:g619424

C:Genetics:

A:Gene: GDB:BMP1; BMP-1

A:Cross-references: GDB:I25203; OMIM:112264

A:Map position: 8p21-8p21

C:Function:

A>Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type I procollagen C-endopeptidase; ascatin homolog; Clr/Cls repeat homology; F:Superfamily: procollagen C-endopeptidase; beta-hydroxyasparagine; bone; calcium; duplication; gl
F:Keywords: alternative splicing; signal sequence #status predicted <SIG>
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-823/Product: procollagen C-endopeptidase splice form HIS #status predicted <MAT>
F:130-321/Domian: astacin homology <AST>
F:322-431/Domian: Clr/Cls repeat homology <CLR1>
F:435-544/Domian: Clr/Cls repeat homology <CLR2>
F:551-587/Domian: EGF homology <EGF>
F:591-700/Domian: Clr/Cls repeat homology <CLR3>
F:738-752/Region: histidine-rich
F:91_142_332_363_599/Binding site: carbohydrate (Asn) #status predicted
F:163-319_185_205_322_348_375-397_435-461_488-510_551-563_559-572_574-587_591-617_644-668
F:213_217_223_272/Binding site: zinc (His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

A; Cross-references: GDB:l125203; OMIM:l12264
A; Map position: 9p21-8p21
C; Function:
C; Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen t
C; Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology;
C; Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication;
F; l-22/Domain: signal sequence #status predicted <SIG>
F; 23-730/Product: procollagen C-endopeptidase splice form BMP1 #status predicted <MAT>
F; l30-321/Domain: astacin homology <AST>
F; 322-431/Domain: Clr/Cls repeat homology <Clr1>
F; 435-544/Domain: Clr/Cls repeat homology <Clr2>
F; 551-587/Domain: EGF homology <EGF>
F; 591-700/Domain: Clr/Cls repeat homology <Clr3>
F; 91-142,332,363,599/Binding site: carboxylate (Asn) (covalent) #status predicted
F; l63-319,185-205,322-348,375-397,435-461,488-510,551-563,559-572,574-587,591-617,644,
F; 213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
F; 214/Active site: Glu #status predicted
F; 565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.9%; Score 183.5; DB 1; Length 730;
Best Local Similarity 38.9%; Pred. No. 2.5e-07;
Matches 49; Conservative 18; Mismatches 44; Indels 15; Gaps 6;

QY 55 NGSTHSKFPHTYPNNMVLVRLVAVDENVRILQTLDERFGLEDPEDDICKYDEVEE- 113
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 599 NGSITSGWPKKEYPPNKCINQLVAPTQ-YRISLOFD---FFETEGNDVKCYDFEVRSG 654
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 114 -PSDGSVLGRWCWGSPVPGKOTSKGNHIRFRVSDEYFPSEPGFCIHYSIIMPVTETTS 172
 : | : | | | | | : | | : | | | | | : | | : | | | | |
Db 655 LTADSKLHGKFCGS-EKPEVITSYNNNRVEFKSDNTV-SKKGFKAHF-----FSEKR 705
 | | : | | | | | : | | : | | | | | : | | : | | | | |
QY 173 PSVLPP 178
 | : | : | |
Db 706 PALQPP 711
 | : | : | |

RESULT 4
JQ0948
A5 antigen precursor - African clawed frog
C; Species: Xenopus laevis (African clawed frog)
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C; Accession: JQ0466; JQ0948
R; Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A; Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homol
A; Reference number: JH0466; MUID:91337458
A; Accession: JH0466
A; Molecule type: mRNA
A; Residues: 1-927 <TAK>
A; Cross-references: GB:D01067; GB:D01077; NID:g222962; PIDN:BARO1260.1; PID:g222963
A; Experimental source: tadpole, brain
A; Note: This protein has motifs homologous to complement components C1r and C1s and
C; Comment: This protein is a neuronal cell surface molecule involved in the neuronal
C; Superfamily: Xenopus A5 antigen; Clr/Cls repeat homology; discoidin I amino-termi
C; Keywords: duplication; glycoprotein; transmembrane protein
F; l-21/Domain: signal sequence #status predicted <SIG>
F; l2-927/Product: A5 antigen #status predicted <ASA>
F; 27-138/Domain: Clr/Cls repeat homology <Clr1>
F; l47-262/Domain: Clr/Cls repeat homology <Clr2>
F; 274-424/Domain: discoidin I amino-terminal homology <DN1>
F; 430-584/Domain: discoidin I amino-terminal homology <DN2>
F; 646-812/Domain: NAM homology <NAM>
F; 861-883/Domain: transmembrane #status predicted <TMW>
F; l50,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 183.5; DB 1; Length 927;
Best Local Similarity 30.7%; Pred. No. 3.4e-07;
Matches 55; Conservative 34; Mismatches 65; Indels 25; Gaps 9;

QY 50 VTISNGSIHSKPFTYPRNMVLVRLVAVDENVRILQTLDFRFGLEDDDDIKYDFV 109
 | : | : | | | | | : | | : | | | | | : | | : | | | | |
 | : | : | | | | | : | | : | | | | | : | | : | | | | |

Db 2233 TSPNHPHYPPHACDIWTLAAPPE-TRIOQFEDRFDIEVTPN--CTSNYLELRDGVSD 2289
Qy 117 GSVLGRWCGSTVPCKQTSKGNHIRIRFVSEYFPEPFCFIHYSI 162
Db 2290 APILSKFCGT-SLPSSQSSGVEVYLFRKSDN-SPTHVGFKA KYSI 2333
RESULT 8
T08618
intrinsic factor-B12 receptor CUBILIN precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08618
R:Moestrup, S.K.; Kosyryaki, R.; Kristiansen, M.; Kaysen, J.H.; Rasmussen, H.H.; Brault, J. Biol. Chem. 273, 5235-5242, 1998
A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies
A:Reference number: Z16459; MUID:98148073
A:Accession: T08618
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-3623 <MOE>
A:Cross-references: EMBL:AF022247; NID:g3834379; PIDN:AAC71661.1; PID:g3834380
C:Genetics:
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membrane
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
F:133-164/Domain: EGF homology <EGF1>
F:436-467/Domain: EGF homology <EGF>
Query Match 8.7%; Score 160; DB 2; Length 3623;
Best Local Similarity 26.8%; Pred. No. 0.00018;
Matches 90; Conservative 44; Mismatches 130; Indels 72; Gaps 23;
Qy 31 QLSDDKEQNGVQDPRHERVVTISGNSIHSKFPFTYPRNMVLYVRLVAVDENVRQLTF 90
Db 924 KFSDDKLECG-----EVLTA-S-TGIITESGHPNVPYRGVNTWHVV-VQRGQLIRLEF 974
Qy 91 DERGLEDEPDICKYDFVEVEEPESDGVLGRWCGSTVPCKQTSKGNHIRIRFVSEYF 150
Db 975 SS-FYLEFHYN--CTNDLEYDTAAOTFLGRYCK-SIPPSLTNSNSIKLIFVSDSAL 1030
Qy 151 PSEPGFCIHY-----SLIMQVVTET-----TSPSVLP---PSS-----LSLDLNNAVT 191
Db 1031 AHE-GFSINTEAIDASSVCLDYDTDFGMLSSPN-FPNYPNSNWEICVIYKITVGLNQOIAL 1088
Qy 192 AFS--TLEELIRYLEPDRQWQVLDLSLYKPTWOLLGKAFLYGKKSKVNVNLLKKEVKLY- 248
Db 1089 HFTDFTLEDYFGSQCVDVEI-RDGGYE-TSPLVG---IY--CGSVLPPTIISHSNKLWL 1141
Qy 249 -----SCTPRNFSVIREELKRTDTIFWPGCLLVKRCGGNACCLHNCNECCQVPRKYT 302
Db 1142 KFKSDAALTAKGSA-----YWDGSG--STGCGGN----LTPTGTGLTSPNYPM 1183
Qy 303 KKYHE---VLQRPKTKVGLKHLKSLTDVALEHHEEC 335
Db 1184 PYHSSECYWRLEASHG-SPPELEFQDFLHHPSC 1218
RESULT 9
A55362
procollagen I C-proteinase enhancer protein precursor - human
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: A55362
R:Takahara, K.; Kessler, E.; Biniaminov, L.; Brusel, M.; Eddy, R.L.; Jani-Sait, S.; Shaw J. Biol. Chem. 269, 26280-26285, 1994
A:Title: Type I procollagen COOH-terminal proteinase enhancer protein: identification, F
A:Reference number: A55362; MUID:95014462
A:Accession: A55362

A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-449 <YAK>
A:Cross-references: GB:L33799; NID:g642907; PIDN:AAA61949.1; PID:g642908
C:Genetics:
A:Gene: GDB:PCOLCE
A:Cross-references: GDB:305468; OMIM:600270
A:Map position: 7q21.3-7q22
C:Superfamily: C1r/C1s repeat homology
C:Keywords: extracellular protein; glycoprotein; pyroglutamic acid
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-449/Product: #status predicted <MAT>
F:37-146/Domain: C1r/C1s repeat homology <C1R1>
F:159-270/Domain: C1r/C1s repeat homology <C1R2>
F:26/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:29,431/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 8.5%; Score 158; DB 2; Length 449;
Best Local Similarity 34.5%; Pred. No. 1.8e-05;
Matches 48; Conservative 21; Mismatches 50; Indels 20; Gaps 9;
Qy 56 GSTHSRPFPHYTPRNMVLYVRLVAVDENVRQLTFDERFGLDEPDICKYDFVEV--- 111
Db 168 GTLTTPNWPESDYPGIGSCSWHIIAPPDQV-IALTFF-EKFDLE--PDYCRYDSVSVEFG 223
Qy 112 EEPDSGVLGRWCGSTVPCKQTSKGNHIRIRFVSEYFPEPFCFIHYSIIM----- 164
Db 224 AVSDDSRLRLKFKCGD-AVPSISSEGNEILLVQFVSDLSVTAD-GFSASYKILPRGTAKEG 281
Qy 165 --PQVTETSPSV-LPFPSS 180
Db 282 QGPGPKRGTEPKVKLPKPS 300
RESULT 10
A39288
dorsal-ventral patterning protein tolloid (EC 3.4.24.-) - fruit fly (Drosophila melan
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39288
R:Shimell, M.J.; Ferguson, E.L.; Childs, S.R.; O'Connor, M.B.
Cell 67, 469-481, 1991
A:Title: The Drosophila dorsal-ventral patterning gene tolloid is related to human bo
A:Reference number: A39288; MUID:92034970
A:Accession: A39288
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1057 <SHI>
A:Cross-references: GB:M76976; NID:g157305; PIDN:AAA28491.1; PID:g157306
C:Genetics:
A:Gene: FlyBase:tlld
A:Cross-references: FlyBase:FBgn0003719
A:Superfamily: dorsal-ventral patterning protein tolloid; astacin homology; C1r/C1s r
C:Keywords: duplication; hydrolase; metalloproteinase; zinc
F:136-329/Domain: astacin homology <AST>
F:352-464/Domain: C1r/C1s repeat homology <C1R1>
F:468-578/Domain: C1r/C1s repeat homology <C1R2>
F:585-620/Domain: EGF homology <EG1>
F:624-740/Domain: C1r/C1s repeat homology <C1R3>
F:747-782/Domain: EGF homology <EG2>
F:787-896/Domain: C1r/C1s repeat homology <C1R4>
F:900-1013/Domain: C1r/C1s repeat homology <C1R5>
F:221,225,231,280/Binding site: zinc (His, His, Tyr) #status predicted
F:222/Active site: Glu #status predicted
Query Match 8.3%; Score 153; DB 1; Length 1057;
Best Local Similarity 33.3%; Pred. No. 0.00014;
Matches 47; Conservative 28; Mismatches 54; Indels 12; Gaps 7;
Qy 13 LAGQRTGTTRAESNLSSKQLQSSDKQNGVQDPRHERV---VTISGNGSIHSRPFPHYTP 68
A:Reference number: A55362; MUID:95014462
A:Accession: A55362

Db 432 VSGEVITQTSRMLLYNVRNAAGYRGFK-ARFEVCGGLKLTQDQSDSPNYPMDYM 490
Qy 69 RNWLVWRLVAVDENVRQLTDFERFGLDEPDDICKYDFVEVEE--PSDGSVLGRWCS 126
Db 491 PDKECVWRITAPD-NHQVAKP-QSFELE--KHDGCAYDFVEIRDGNHSDSRLLGRFCGD 546
Qy 127 GTVPGKQTSKGNHIRRFVSD 147
Db 547 KLPPNIKT-RSNQMYIRFVSD 566
RESULT 11
T30337
Polyprotein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30337
R:Yang, J.C.; Lindsay, L.L.; Hedrick, J.L.
submitted to the EMBL Data Library, March 1998
A:Description: CDNA cloning of ovocytinase, a chymotrypsin-like protease released from X.
A:Reference number: Z20829
A:Accession: T30337
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-1524 <YAN>
A:Cross-references: EMBL:U01290; NID:g2981640; PID:g2981641; PIDN:AAC24717.1
Query Match 8.0%; Score 147.5; DB 2; Length 1524;
Best Local Similarity 28.7%; Pred. No. 0.00066;
Matches 54; Conservative 34; Mismatches 73; Indels 27; Gaps 10;
Qy 56 GSIHSPKPFHTYPRNNVLWRLVAVDENVRQLTDFERFGLDEPDDICKYDFVEV-EEP 114
Db 439 GMIYSPNPDPYPRUKTCSW-IIEAPENHIVLKFED-ENVEYGHG-CIYDAVEYDGA 494
Qy 115 SDGSVLGRWCGSGTVPGKQTSKGNHIRRFVSDYFPSPGFCIHYSIMQP-----VT 168
Db 495 EKKQIARLCGY-TLPLTSPSPENTMLIRFKTD-MENSYPGFKVFKFSPKQFSLPVD 552
Qy 169 EFTSPSVLPSPSLDNLNA-VTAFSTLEELIRYLEPD-----RWQVDL-----DS 214
Db 553 DPTTSMHLPRALDVCGMAPMTPKWLPRIYVGGEASPNWPQVQVIFLRTFCEGA 612
Qy 215 LYKPTWQL 222
Db 613 IISPOWIL 620
RESULT 12
A59271
Ra-reactive factor (EC 3.4.21.-) 2 precursor - human
N:Alternate names: mannose binding protein-associated serine proteinase 2 (MASP-2)
C:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 16-Jun-2000
C:Accession: A59271
R:Thiel, S.; Vorup-Jensen, T.; Stover, C.M.; Schwaebler, W.J.; Laursen, S.B.; Poulsen, K.
Nature 386, 506-510, 1997
A:Title: A second serine protease associated with mannan-binding lectin that activates
A:Reference number: A59271; MUID:97242412
A:Accession: A59271
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-686 <JEN>
A:Cross-references: GB:Y09926; NID:g4007626; PIDN:CRA71059.1; PID:g4007627
A:Experimental source: tissue liver
A:Note: submitted to Genbank, December 1996
A:Note: parts of this sequence, including the amino end of the mature protein, were deter
C:Genetics:
A:Gene: GDB:MASP2
A:Cross-references: GDB:6071500
A:Map position: lp36.2-lp36.3
C:Superfamily: complement subcomponent C1r; C1r/C1s repeat homology; complement factor H

C:Keywords: beta-hydroxyasparagine; complement pathway; duplication; hydrolase; serin
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-444,445-686/Product: Ra-reactive factor 2 #status predicted <WAT>
F:19-134/Domain: C1r/C1s repeat homology <C1R1>
F:142-180/Domain: EGF homology <EGF>
F:184-293/Domain: C1r/C1s repeat homology <C1R2>
F:300-361/Domain: complement factor H repeat homology <FHL>
F:366-430/Domain: complement factor H repeat homology <FH2>
F:445-679/Domain: trypsin homology <TRY>
F:72-90,142-156,152-165,167-180,184-211,241-259,300-348,328-361,366-412,396-430,434-5
F:158/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:444-445/Cleavage site: Arg-Ile (autolytic) #status predicted
F:483,532,633/Active site: His, Asp, Ser #status predicted

Query Match 7.9%; Score 145.5; DB 1; Length 686;
Best Local Similarity 30.8%; Pred. No. 0.00035;
Matches 33; Conservative 27; Mismatches 42; Indels 5; Gaps 4;
Qy 55 NGSIHSPKPFHTYPRNNVLWRLVAVDENVRQLTDFERFGLDEPDDICKYDFVEVEEP 114
Db 193 SGELSSPEYPRPYKLSSTYS-ISLEEGFSVILDFVESFDVTHPETLCPYDLKIQ-- 249
Qy 115 SDGSVLGRWCGSGTVPGKQTSKGNHIRRFVSDYFPSPGFCIHYS 161
Db 250 TDREHGFPGCK-TLPRIETKSNVTITFTVDE-SGDHTGKWKIHYT 294

RESULT 13
I54763
Ra-reactive factor (EC 3.4.21.-) 1 precursor - human
N:Alternate names: mannose binding protein-associated serine proteinase 1 (MASP-1)
C:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 16-Jun-2000
C:Accession: I54763; JN0883
R:Sato, T.; Endo, Y.; Matsushita, M.; Fujita, T.
Int. Immunol. 6, 665-669, 1994
A:Title: Molecular characterization of a novel serine protease involved in activation
A:Reference number: I54763; MUID:94289349
A:Accession: I54763
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-699 <SAT>
A:Cross-references: GB:D2593; NID:g790963; PIDN:BAA05928.1; PID:g471128
R:Takada, F.; Takayama, Y.; Hattuse, H.; Kawakami, M.
Biochem. Biophys. Res. Commun. 196, 1003-1009, 1993
A:Title: A new member of the C1s family of complement proteins found in a bactericida
A:Reference number: JN0883; MUID:94059062
A:Accession: JN0883
A:Molecule type: mRNA
A:Residues: 1-234, 'E', 236-284, 'G', 286-498, 'K', 500-542, 'K', 544-642, 'S', 644-699 <TAK>
A:Cross-references: DDBJ:D17525; NID:g439712; PIDN:BAA04477.1; PID:g439713
A:Experimental source: liver
C:Comment: This is a serum bactericidal factor that activates complement C4 and C2 co

C:Genetics:
A:Gene: GDB:MASP1; GDB:CRARF1; PRSS5; MASP
A:Cross-references: GDB:361104; GDB:330954; OMIM:600521
A:Map position: 3q27-3q28
C:Superfamily: complement subcomponent C1r; C1r/C1s repeat homology; complement facto
C:Keywords: beta-hydroxyasparagine; complement pathway; duplication; glycoprotein; hy
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-448,449-699/Product: Ra-reactive factor #status predicted <WAT>
F:19-135/Domain: C1r/C1s repeat homology <C1R1>
F:143-181/Domain: EGF homology <EGF>
F:185-294/Domain: C1r/C1s repeat homology <C1R2>
F:300-362/Domain: complement factor H repeat homology <FHL>
F:367-432/Domain: complement factor H repeat homology <FH2>
F:449-691/Domain: trypsin homology <TRY>
F:49,178,407/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:73-91,143-157,153-166,168-181,185-212,242-260,301-349,329-362,367-414,397-432,436-5
F:159/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:448-449/Cleavage site: Arg-Ile (autolytic) #status predicted
F:490,552,646/Active site: His, Asp, Ser #status predicted

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:03 ; Search time 12.6 Seconds
(without alignments)
937.948 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 1848
Sequence: 1 MLLGLLLLTSLAGQRTGT.....DVALEHHEEDCVCRGNAGG 345

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	191	10.3	707	1	BMPL_XENLA	P98070 xenopus lae
2	183.5	9.9	928	1	NRPL_XENLA	P28924 xenopus lae
3	181	9.8	986	1	BMPL_HUMAN	P13497 homo sapien
4	181	9.8	991	1	BMPL_MOUSE	P98063 mus musculus
5	176	9.5	923	1	NRPL_HUMAN	O14786 homo sapien
6	172	9.3	922	1	NRPL_RAT	O9qwj9 rattus norv
7	169	9.1	923	1	NRPL_MOUSE	P97333 mus musculus
8	167	9.0	914	1	NRPL_CHICK	P79795 gallus gall
9	163.5	8.8	616	1	SPAN_STRPU	P98068 strongyloce
10	158	8.5	449	1	PCOL_HUMAN	Q15113 homo sapien
11	155.5	8.4	925	1	NRPL_RAT	O35276 rattus norv
12	155.5	8.4	931	1	NRPL_HUMAN	O60462 homo sapien
13	155.5	8.4	931	1	NRPL_MOUSE	O35375 mus musculus
14	153	8.3	1057	1	TLD_DROME	P25723 drosophila
15	147.5	8.0	704	1	CRAR_MOUSE	P98064 mus musculus
16	145.5	7.9	686	1	MAS2_HUMAN	O00187 homo sapien
17	144.5	7.8	597	1	BP10_PARLI	P42674 paracentrot
18	144	7.8	699	1	CRAR_HUMAN	P48740 h complemen
19	137.5	7.4	705	1	CLR_HUMAN	P00736 homo sapien
20	136.5	7.4	468	1	PCOL_RAT	O08628 rattus norv
21	133.5	7.2	468	1	PCOL_MOUSE	O61398 mus musculus
22	132	7.1	415	1	VEGC_MOUSE	P97953 mus musculus
23	130	7.0	639	1	BMPL_STRPU	P98069 strongyloce
24	128	6.9	277	1	TSG6_HUMAN	P98066 homo sapien
25	127	6.9	419	1	VEGC_HUMAN	P49767 homo sapien
26	125.5	6.8	245	1	PDGB_FELCA	P12919 felis silve
27	117.5	6.4	276	1	TSG6_RABIT	P98065 oryctolagus
28	114.5	6.2	226	1	PDGA_XENLA	P13698 xenopus lae
29	113	6.1	695	1	CASP_MESAU	P15156 mesocricetu
30	112.5	6.1	241	1	PDGB_HUMAN	P01127 homo sapien
31	111.5	6.0	164	1	VEGF_CAVPO	P26617 cavia porce.
32	111.5	6.0	226	1	TSIS_SMSAV	P01128 simian sarc
33	109	5.9	855	1	ST14_HUMAN	O9y5y6 homo sapien

ALIGNMENTS

```
RESULT 1
BMPL_XENLA
ID BMPL_XENLA STANDARD; PRT; 707 AA.
AC P98070;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.-) (BMP-1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94085787; PubMed=8262384;
RA Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;
RT "Cloning and expression of cDNA encoding Xenopus laevis bone
RT morphogenetic protein-1 during early embryonic development.";
RL Gene 134:257-261(1993).
CC -1- FUNCTION: INVOLVED IN PATTERN FORMATION IN GASTRULA AND LATER
CC DIFFERENTIAL STAGE OF DEVELOPING ORGANS.
CC -1- DEVELOPMENTAL STAGE: BLASTULA, EARLY GASTRULA AND HATCHED
CC TADPOLES; LITTLE OR NO EXPRESSION IN MORULA AND LATE GASTRULA.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 CUB DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L12249; AAA16313.1;
CC HSSP; P00736; IAPQ.
CC MEROPS; M12.005;
CC InterPro; IPR000130;
CC InterPro; IPR000152;
CC InterPro; IPR000561;
CC InterPro; IPR000859;
CC InterPro; IPR001506;
CC InterPro; IPR001881;
CC Pfam; PF01400; Astacin; 1.
CC Pfam; PF00431; Astacin; 3.
CC Pfam; PF00008; EGF; 1.
CC PRINTS; PR00480; ASTACIN.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS01180; CUB; 3.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
```

DR PROSITE: PS01187; EGF_CA; 1.
 KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase; Protease;
 KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP 83 POTENTIAL.
 FT CHAIN 84 707 BONE MORPHOGENETIC PROTEIN 1.
 FT DOMAIN 84 284 METALLOPROTEASE.
 FT DOMAIN 285 397 CUB.
 FT DOMAIN 398 509 CUB.
 FT DOMAIN 510 551 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 554 666 CUB.
 FT METAL 176 177 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 177 177 BY SIMILARITY.
 FT METAL 180 180 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 186 186 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 514 526 BY SIMILARITY.
 FT DISULFID 522 535 BY SIMILARITY.
 FT DISULFID 537 550 BY SIMILARITY.
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC9B8D CRC64;
 Query Match 10.3%; Score 191; DB 1; Length 707;
 Best Local Similarity 45.2%; Pred. No. 3.2e-08;
 Matches 4; Conservative 16; Mismatches 33; Indels 8; Gaps 5;
 QY 55 NGSINSPGFHTYPRNMVLRVAVDENVRQITFDERFGLDEPDICKYDFVEVEE- 113
 DB 562 NGSINSPGWPEYPPNKKCIQLVAPTO-YRISLKFDQ---FETGNDVCKYDFVEVRSG 617
 QY 114 -PSDGSVLGRWCGSTVPGKTSKGNHRIKRVSDYFPSPGPF 156
 DB 618 LTSDSKLHGKFCGS-ELPAVITTSQNNMRIBFKSDNTV-SKKG 659.
 RESULT 2
 NRPL_XENLA
 ID NRPL_XENLA STANDARD; PRT; 928 AA.
 AC P28824;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN) (A5 ANTIGEN).
 OC Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX Medline=9137458; Pubmed=1908252;
 RA Takagi S., Hirata T., Agata K., Mochii M., Eguchi G., Fujisawa H.;
 RT "The A5 antigen, a candidate for the neuronal recognition molecule,
 RT has homologues to complement components and coagulation factors.";
 RL Neuron 7:295-307(1991).
 CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
 CC CARDIOVASCULAR SYSTEM. IN ANGIOGENESIS, IN THE FORMATION OF
 CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
 CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
 CC SIMILARITY). PRESUMED TO BE INVOLVED IN THE NEURONAL RECOGNITION
 CC BETWEEN THE OPTIC NERVE FIBERS AND THE VISUAL CENTERS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: RETINAL GANGLION CELLS AND VISUAL CENTER
 CC NEURONS.
 CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL: D10467; BAA01260.1; -
 DR InterPro: IPR000421; -
 DR InterPro: IPR000859; -
 DR InterPro: IPR000998; -
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR Pfam: PF00629; MAM; 1.
 DR PRINTS: PR00020; MAMDOMAIN.
 DR PROSITE: PS00740; MAM_1; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS50060; MAM_2; 1.
 KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Antigen.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 928 NEUROFILIN-1.
 FT DOMAIN 22 860 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 861 883 POTENTIAL.
 FT DOMAIN 884 928 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 27 141 CUB 1.
 FT DOMAIN 147 265 CUB 2.
 FT DOMAIN 275 424 F5/8 TYPE C 1.
 FT DOMAIN 431 584 F5/8 TYPE C 2.
 FT DOMAIN 646 812 MAM.
 FT DISULFID 27 54 PROBABLE.
 FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT DISULFID 431 584 BY SIMILARITY.
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 928 AA; 103416 MW; AF6B323B0A4C789D CRC64;
 Query Match 9.9%; Score 183.5; DB 1; Length 928;
 Best Local Similarity 30.7%; Pred. No. 1.9e-07;
 Matches 55; Conservative 34; Mismatches 65; Indels 25; Gaps 9;
 QY 50 VTISNGSIHSPKPHPTPRNMVLRVAVDENVRQITFDERFGLDEPDICKYDFV 109
 DB 31 IKITSPSYLTSGYSPHSPQCEWLIQAPEHYQRIMINFNPHFLEDE---CKYDYV 87
 QY 110 EV--EPPSDGSVLGRWCGSTVPGKTSKGNHRIKRVSDYFPSPGFCIHYSIIM--P 165
 DB 88 EVIDGNANGQLLGKYGCGK-IAPSPVLTGSPSIFIRFVSDYETPG-AGFSIRVEVFTGP 145
 QY 166 QVTE--TTSFSPVLP-----PSSLSLDLNNVATFSTLEELIRYLEPDRQVLDLS 214
 DB 146 ECSRNETSSNGVIKSPKYPEKYPNALECYIIFA----PKMQEIV--LFESELEADS 198
 RESULT 3
 BMPL_HUMAN
 ID BMPL_HUMAN STANDARD; PRT; 986 AA.
 AC P13497; Q13292; Q99421; Q99422; Q99423; Q14874;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DR	Pfam; PF01400; Astacin; 1.
DR	Pfam; PF00431; CUB; 5.
DR	Pfam; PF00008; EGF; 2.
DR	PRINTS; PR00480; ASTACIN.
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.
DR	PROSITE; PS01180; CUB; 5.
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.
DR	PROSITE; PS01186; EGF_2; 2.
DR	PROSITE; PS01187; EGF_CA; 2.
KW	Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase;
KW	Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
KW	Glycoprotein; Zymogen.
FT	SIGNAL 1 25
FT	PROPEP 26 125
FT	CHAIN 126 991
FT	DOMAIN 126 326
FT	DOMAIN 327 439
FT	DOMAIN 440 551
FT	DOMAIN 552 593
FT	DOMAIN 596 707
FT	DOMAIN 708 748
FT	DOMAIN 752 864
FT	DOMAIN 865 981
FT	METAL 218 218
FT	ACT_SITE 219 219
FT	METAL 222 222
FT	METAL 228 228
FT	DISULFID 327 353
FT	DISULFID 380 402
FT	DISULFID 440 466
FT	DISULFID 493 515
FT	DISULFID 556 568
FT	DISULFID 564 577
FT	DISULEID 579 592
FT	DISULFID 596 622
FT	DISULFID 649 671
FT	DISULFID 712 723
FT	DISULFID 719 732
FT	DISULFID 734 747
FT	CARBOHYD 96 96
FT	CARBOHYD 147 147
FT	CARBOHYD 337 337
FT	CARBOHYD 368 368
FT	CARBOHYD 604 604
SEQ	SEQUENCE 991 AA; 111607 MW; 68A1847783A0BB9E CRC64;
Query Match 9.8%; Score 181; DB 1; Length 991;	
Best Local Similarity 42.6%; Pred.No.3.e-07;	
Matches 46; Conservative 16; Mismatches 38; Indels 8; Gaps	
QY	55 NGSIHSPKFPHTYPRNMLVRLVAVDENVRIOLTDEFRGLGEDPDCKYDFVEEE- 113 : : : : : :
Db	604 NGSIISPWPKEYPNNKNKIQLVAPTQ-YRISLOFD---FFETEGNDVKYDFVEVRSG 659
QY	114 -PDGSLVGRWCSSGTVPGRQTSGKHNIIRIVSYDEVFSEPGFCIHY 160 : : : : : : : :
Db	660 LTADSKLHRFCGS-EKPEVITSOYNNMRVFESDNVT-SKRGFKAHF 705
RESULT 5	
ID	NRP1_HUMAN STANDARD; PRT; 923 AA.
AC	O14786; O60461;
DT	30-MAY-2000 (Rel. 39, Created)
DI	01-OCT-2000 (Rel. 40, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	NEUROFILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
GN	NRP1 OR NRP OR VEGF165R.
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBITaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
RX MEDLINE=97433084; PubMed=9288753;
RA He Z., Tessier-Lavigne M.;
RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin
RT III,";
RL Cell 90:739-751(1997).
RN [2]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
RC TISSUE=Brain;
RX MEDLINE=9818099; PubMed=9529250;
RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor,";
RL Cell 92:735-745(1998).
RN [3]
RP SEQUENCE FROM N.A. (SOLUBLE/SNRP1 ISOFORM), AND SEQUENCE OF 22-31.
RC TISSUE=Prostatic adenocarcinoma;
RX MEDLINE=2018929; PubMed=1068880;
RA Gagnon M.L., Bielenberg D.R., Gerechtman Z., Miao H.-Q., Takashima S.,
RA Soker S., Klagsbrun M.;
RT "Identification of a natural soluble neuropilin-1 that binds vascular
RT endothelial growth factor: In vivo expression and antitumor
RT activity,";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20309748; PubMed=10748121;
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
RT form of vascular endothelial growth factor (VEGF) and of placenta
RT growth factor-2, but only neuropilin-2 functions as a receptor for
RT the 145-amino acid form of VEGF,";
RL J. Biol. Chem. 275:18040-18045(2000).
RN [5]
RP FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE
RC DEVELOPMENT OF THE CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE
CC FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS
CC OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOPULSANT
CC ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2
CC ISOFORM OF PGF, THE VEGF-165 ISOFORM OF VEGF AND VEGF-B.
CC COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR
CC AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED
CC ANGIOGENESIS.
CC FUNCTION: THE SOLUBLE/SNRP1 ISOFORM BINDS VEGF-165 AND APPEARS
CC TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY
CC SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE
CC SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD
CC VESSEL NUMBER AND INTEGRITY.
CC SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRP1
CC ISOFORM IS SECRETED.
CC ALTERNATIVE PRODUCTS: 2 ISOFORMS; MEMBRANE-BOUND (SHOWN HERE)
CC AND SOLUBLE/SNRP1; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND
CC SOLUBLE/SNRP1 ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-
CC BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT
CC TISSUES. IN THE DEVELOPING EMBRYO IT IS FOUND PREDOMINANTLY IN THE
CC NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART
CC AND PANCREAS; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY
CC AND PLACENTA; AND LOW IN ADULT BRAIN. THE SOLUBLE/SNRP1 ISOFORM IS
CC FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
CC SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AF018956; AAC51759.1; -;
CC EMBL; AF016050; AAC12921.1; -;
CC EMBL; AF145712; AAF44344.1; -;
CC MIM; 602069; -;
CC InterPro: IPR000421; -;
CC InterPro: IPR000859; -;
CC InterPro: IPR000998; -;
CC Pfam; PF00431; CUB; 2;
CC Pfam; PF00754; F5_F8_type_C; 2;
CC Pfam; PF00629; MAM; 1;
CC PRINTS; PR00020; MAMDOMAIN.
CC PROSITE; PS01180; CUB; 2;
CC PROSITE; PS01285; FAS8C_1; 2;
CC PROSITE; PS01286; FAS8C_2; 2;
CC PROSITE; PS00740; MAM_1; 1;
CC PROSITE; PS00060; MAM_2; 1;
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 22 923 NEUROPILIN-1.
FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 857 879 POTENTIAL.
FT DOMAIN 880 923 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 642 644 EFP -> GIK (IN SOLUBLE/SNRP1 ISOFORM).
FT VARSPLIC 645 923 MISSING (IN SOLUBLE/SNRP1 ISOFORM).
FT CONFLICT 26 26 K -> E (IN REF. 1).
FT CONFLICT 749 749 D -> H (IN REF. 2).
FT CONFLICT 855 855 E -> D (IN REF. 2).
SQ SEQUENCE 923 AA; 103120 MW; ADEADC4A849E5D57 CRC64;

Query Match 9.5%; Score 176; DB 1; Length 923;
Best Local Similarity 32.6%; Pred. No. 7.8e-07;
Matches 57; Conservative 24; Mismatches 68; Indels 26; Gaps 9;

QY 50 VTISGNSIHSKPHPTPHYPRMVLVRLVAVDENVRQLTDFERFGLDEPDICKYDFV 109
DB 31 IKTESGYLSPGYPHSYHPSEKCEWLIQADPPYQRMNINPNHFDLEDRD---CKYD 87
QY 110 EV--EEPSDGSVLGRWCGSGTVPCQKSGNHIRFVSDVFPSEPGFCIHSIIM--P 165
DB 88 EVDGNGENHGRKFCGK-IAPPVVSNGPFLFKVSD-YETHGAGFSRYEIKRGP 145
QY 166 QVTE--TTSPSVLP-----PSSLSDLLNNVATFSTLEELIRY----LEPD 206
DB 146 ECSQNYTTPSGVKSFGPEKYPNSLECTYI---VFAPKMSIILEFSDLEPD 197

RESULT 6
NRPI_RAT ID NRPI_RAT STANDARD; PRT; 922 AA.
AC Q9QWJ9; 2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)

	* NEUROPILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
DE	
EN	NRP1.
GN	Rattus norvegicus (Rat).
OS	Oc Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX	NCBI_TaxID=10116;
ON	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=SPRAGUE-DAWLEY;
RF	MEDLINE=97433085; PubMed=9288754;
RX	Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J.,
RA	Ginty D.D.;
RD	"Neuropilin is a semaphorin III receptor.";
RE	Cell 90:753-762(1997).
RT	-1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC	CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC	CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
CC	SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. IT
CC	BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165
CC	ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN
CC	INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS.
CC	IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY).
CC	SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC	-1- TISSUE SPECIFICITY: FOUND IN THE EMBRYONIC NERVOUS SYSTEM.
CC	-1- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
CC	-1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@lsb-sib.ch).
CC	-----
CC	EMBL; AF016296; AAC53337.1; ..
DR	InterPro: IPRO00421; ..
DR	InterPro: IPRO00859; ..
DR	InterPro: IPRO00998; ..
DR	Pfam; PF000431; CUB; 2.
DR	Pfam; PF00629; MAM; 1.
DR	Pfam; PF00754; F5_F8_type_C; 2.
DR	PRINTS: PR00020; MAMDMAIN.
DR	PROSITE; PS01180; CUB; 2.
DR	PROSITE; PS01285; FA58C_1; 2.
DR	PROSITE; PS01286; FA58C_2; 2.
DR	PROSITE; PS00740; MAM.1; 1.
DR	PROSITE; PS50060; MAM.2; 1.
KW	Transmembrane; Glycoprotein; Neuregulin; Signal; Repeat; Receptor.
FT	SIGNAL 1 21 POTENTIAL.
FT	CHAIN 22 922 NEUROPILIN-1.
FT	DOMAIN 22 855 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 856 880 POTENTIAL.
FT	DOMAIN 881 922 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 27 141 CUB 1.
FT	DOMAIN 147 265 CUB 2.
FT	DOMAIN 275 424 F5/F8 TYPE C 1.
FT	DOMAIN 431 583 F5/F8 TYPE C 2.
FT	DOMAIN 645 811 MAM.
FT	DISELFD 27 54 PROBABLE.
FT	DISELFD 82 104 PROBABLE.
FT	DISELFD 147 173 PROBABLE.
FT	DISELFD 206 228 PROBABLE.
FT	DISELFD 275 424 BY SIMILARITY.
FT	DISELFD 431 583 BY SIMILARITY.
FT	CARBOHYD 150 150 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD 261 261 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD 300 300 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD 522 522 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD 841 841 N-LINKED (GLCNAC..) (POTENTIAL).

```

DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 1.
KW Transmembrane; Glycoprotein;
FT SIGNAL 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 583
FT DOMAIN 645 811
FT DISULFID 27 54
FT DISULFID 82 104
FT DISULFID 147 173
FT DISULFID 206 228
FT DISULFID 275 424
FT DISULFID 431 583
FT CARBOHYD 150 150
FT CARBOHYD 261 261
FT CARBOHYD 300 300
FT CARBOHYD 522 522
FT CARBOHYD 842 842
SQ SEQUENCE 923 AA; 103020 MW; 0644B8A170796808 CRC64;

Query Match 9.1%; Score 169; DB 1; Length 923;
Best Local Similarity 31.5%; Pred. No. 2.9e-06;
Matches 56; Conservative 21; Mismatches 67; Indels 34; Gaps 8;

QY 2 LLGLLLTSLALAGQTGTAEBSNLSSKLQLSDREQNGVQDPRHVVVTISGNGSIHSP 61
DB 7 LLCATLALALAG-----AFRSDKCG-----TKIENPGVLTSP 42
QY 62 KPHPTPRMVLVRLVAVDENVRQLTDFERFGLDEPDDICKYDFVEV--EPPSDGVS 119
DB 43 GYPHSYHPSKCEWLTAQEPYQRIIINFNFHFDLEDRD---CKDYVEVDIGENEGRL 99
QY 120 LGRWCGSGTVPGKQTSKGNHRIIRVSDYFSEPGFCIHYSIIM--PQVTEI-TSPS 174
DB 100 WKFCGK-IAPSPVYSSGPFLLFIKTVSD-YETHGAGFSIRYIFKRGPCSQNYTAP 155

RESULT 8
ID NRPI_CHICK STANDARD; PRT; 914 AA.
AC P79795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
GN NRPI OR NRP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE=Embryonic brain;
RX MEDLINE=95324761; PubMed=7601310;
RA Takagi S., Kesuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
RA Fujisawa H.;
RT "Expression of a cell adhesion molecule, neurofilin, in the
RL developing chick nervous system.";
RL Dev. Biol. 170:207-222(1995).
CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY

```

```

CC SIMILARITY). SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
CC PROPERTIES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM; OPTIC TECTUM
CC (LAYERS D AND E OF SGFS), AMACRINE CELLS OF RETINA, NEURITES OF
CC DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
CC BLOOD VESSELS IN THE ENTIRE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D45416; BAA08256.1; -.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00629; MAM; 1.
CC PRINTS; PR00020; MAMDOMAIN.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS00060; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Cell adhesion.
KW SIGNAL 1 18
KW CHAIN 19 914
FT DOMAIN 20 847
FT TRANSMEM 848 870
FT DOMAIN 871 914
FT DOMAIN 25 139
FT DOMAIN 145 263
FT DOMAIN 273 422
FT DOMAIN 429 581
FT DOMAIN 636 801
FT DISULFID 25 52
FT DISULFID 80 102
FT DISULFID 145 171
FT DISULFID 204 226
FT DISULFID 273 422
FT DISULFID 429 581
SQ SEQUENCE 914 AA; 102480 MW; DD2EE6D6F0CBB68C CRC64;

Query Match 9.0%; Score 167; DB 1; Length 914;
Best Local Similarity 28.8%; Pred. No. 4.3e-06;
Matches 51; Conservative 31; Mismatches 65; Indels 30; Gaps 9;

QY 50 VTISGNGSIHSPKPHPTPRMVLVRLVAVDENVRQLTDFERFGLDEPDDICKYDFV 109
DB 29 IKILSPGYLTSPGYPSQYHPSQKCEWLTAQEPYQRIIINFNFHFDLEDRD---CKDYV 85
QY 110 EV--EPPSDGSGTVPGKQTSKGNHRIIRVSDYFSEPGFCIHYSIIM--P 165
DB 86 EVIDGDNAEGRLMGKCGK-IAPPLVSSGYPYLFKTVSD-YETHGAGFSIRYEVFRGP 143
QY 166 QVTE--TTPSPVLP-----PSSLSDL-----LNNAVTAFSTLEELIRYLEPD 206
DB 144 ECSRNFTSSSGMIKSPGPKYKPNSECTYIIFAPKMSIILEFESFE----LEPD 195

RESULT 9
SPAN_STRPU

```


[illegible]

Query Match 8.4%; Score 155.5; DB 1; Length 925;
 Best Local Similarity 33.3%; Pred. No. 3.9e-05;
 Matches 41; Conservative 19; Mismatches 56; Indels 7; Gaps 4;

QY 42 QDPHERVVTISGNGSIHSPKPHPTYPNWLVLVAVDENVRQLTDFEGLDEPED 101
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 24 QDPCCGRLNKDAGYITSPGYPDYSHQNCWEVVIYAPENQKIVLNFPHFEIEKHD- 82
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 QY 102 DICKYDFVEVEEPSDGS--VLGRWCGSGTVPGKQTSKGNHIRFVSDYFSEPGFCIH 159
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 83 --CKYDFIEIRDGSESADLLGKHGCGN-IAPPTIISGSLVLYIKFTSD-YARQAGFSLR 138
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 QY 160 YSI 162
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 139 YEI 141
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

RESULT 12
 NRP2_HUMAN STANDARD; PRT; 931 AA.
 ID NRP2_HUMAN STANDARD; PRT; 931 AA.
 AC O60462; O14820; O14821;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROPILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).
 DE NEUROPILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).
 GN NRP2 OR VEGF165R2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=97470888; PubMed=9331348;
 RX Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
 RA "Neuropilin-2, a novel member of the neuropilin family, is a high
 RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
 RL III.";
 RL Neuron 19:547-559(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A0 AND A17).
 RX MEDLINE=98188099; PubMed=9529250;
 RA Soker S., Takashima S., Miao H.-O., Neufeld G., Klagsbrun M.;
 RT "Neuropilin-1 is expressed by endothelial and tumor cells as an
 RL isoform-specific receptor for vascular endothelial growth factor.";
 RL Cell 92:735-745(1998).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20309748; PubMed=10748121;
 RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
 RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
 RT form of vascular endothelial growth factor (VEGF) and of placenta
 RT growth factor-2, but only neuropilin-2 functions as a receptor for
 RT the 145-amino acid form of VEGF.";
 RL J. Biol. Chem. 275:18040-18045(2000).
 CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165
 CC AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.
 CC -1- SUBUNIT: NEUROPILIN-2 PROBABLY FORMS AN HETEROMERIC COMPLEX WITH
 CC NEUROPILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN 3C RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A22 (SHOWN HERE), A0 AND A17;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; AF022859; AAC51788.1; -
 CC EMBL; AF022860; AAC51789.1; -
 CC EMBL; AF016098; AAC12922.1; -
 CC MIM; 602070; -
 DR InterPro; IPR000421; -
 DR InterPro; IPR000859; -
 DR InterPro; IPR000998; -
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00629; MAM; 1.
 DR Prosite; PS01180; F5_F8_type_C; 2.
 DR Prosite; PS01285; FA58C.1; 2.
 DR Prosite; PS01286; FA58C.2; 2.
 DR Prosite; PS01286; MAM_2; 1.
 KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Alternative splicing.
 FT SIGNAL 1 20 OR 22 (POTENTIAL).
 FT CHAIN 21 931 NEUROPILIN-2.
 FT DOMAIN 21 864 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 865 889 POTENTIAL.
 FT DOMAIN 890 931 CYTOPLASMIC (POTENTIAL).
 FT CUB 1. CUB 1.
 FT CUB 2. CUB 2.
 FT F5/8 TYPE C 1. F5/8 TYPE C 1.
 FT F5/8 TYPE C 2. F5/8 TYPE C 2.
 FT MAM. MAM.
 FT POLY-SER. POLY-SER.
 FT BY SIMILARITY. BY SIMILARITY.
 FT BY SIMILARITY. BY SIMILARITY.
 FT BY SIMILARITY. BY SIMILARITY.
 FT BY SIMILARITY. BY SIMILARITY.
 FT BY SIMILARITY. BY SIMILARITY.
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT MISSING (IN ISOFORM A17).
 FT MISSING (IN ISOFORM A0).
 FT E -> K (IN REF. 1).
 SQ SEQUENCE 931 AA; 104830 MW; 270CBAE69A0A797C CRC64;
 Query Match 8.4%; Score 155.5; DB 1; Length 931;
 Best Local Similarity 32.1%; Pred. No. 3.9e-05;
 Matches 42; Conservative 20; Mismatches 62; Indels 7; Gaps 4;

QY 34 SKEQNGVDPHERVVTISGNGSIHSPKPHPTYPNWLVLVAVDENVRQLTDFER 93
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 16 SRHQVRGQDPGCGRLNKDAGYITSPGYPDYSHQNCWEVVIYAPENQKIVLNFPH 75
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 QY 94 FGLEDDEDDICKYDFVEVEEPSDGS--VLGRWCGSGTVPGKQTSKGNHIRFVSDYF 151
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 76 FEIEKHD---CKYDFIEIRDGSESADLLGKHGCGN-IAPPTIISGSLVLYIKFTSD-YAR 130
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 QY 152 SEPGFCIHYSI 162
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
 Db 131 QGAGFSLRVEI 141
 ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

RESULT 13
 NRP2_MOUSE STANDARD; PRT; 931 AA.
 ID NRP2_MOUSE STANDARD; PRT; 931 AA.
 AC O35375; O35373; O35374; O35376; O35377; O35378;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROPILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).

ID AC CRAR_MOUSE STANDARD; PRT; 704 AA.
 DT P98064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COMPLEMENT-ACTIVATING COMPONENT OF RA-REACTIVE FACTOR PRECURSOR
 DE (EC 3.4.21.-) (RA-REACTIVE FACTOR SERINE PROTEASE P100) (RARE)
 DE (MANNAN-BINDING LECTIN SERINE PROTEASE 1).
 DE MASPI OR CRARF.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=94179811; PubMed=8133044;
 RA Takayama Y., Takada F., Takahashi A., Kawakami M.;
 RT "A 100-kDa protein in the C4-activating component of Ra-reactive
 RT factor is a new serine protease having module organization similar to
 RT C1r and C1s.";
 RL J. Immunol. 152:2308-2316(1994).
 RN [2]
 RP SEQUENCE OF 465-704 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=93176166; PubMed=8439319;
 RA Takahashi A., Takayama Y., Hattuse H., Kawakami M.;
 RT "Presence of a serine protease in the complement-activating component
 RT of the complement-dependent bactericidal factor, RaRF, in mouse
 RT serum.";
 RL Biochem. Biophys. Res. Commun. 190:681-687(1993).
 CC -!- FUNCTION: COMPONENT OF THE BACTERICIDAL RA-REACTIVE FACTOR RARF
 CC WHICH SPECIFICALLY BINDS TO RA AND R2 POLYSACCHARIDES EXPRESSED BY
 CC CERTAIN ENTEROBACTERIA. IT TRIGGERS THE ACTIVATION OF COMPLEMENT
 CC CASCADE BY ACTIVATING THE C4 AND C2 COMPONENTS. IT ACTIVATES THE
 CC C4 COMPONENT BY CLEAVING THE ALPHA-CHAIN OF C4.
 CC -!- SUBUNIT: RARF CONSISTS OF A COMPLEMENT-ACTIVATING COMPONENT
 CC (CRARF) AND A POLYSACCHARIDE-BINDING (MANNOSE-BINDING) COMPONENT.
 CC CRARF IS AN HETERODIMER OF A HEAVY (P70) AND A LIGHT CHAIN (29)
 CC LINKED BY A DISULFIDE BOND.
 CC -!- TISSUE SPECIFICITY: LIVER.
 CC -!- DOMAIN: CRARF HAS A MODULE ORGANIZATION SIMILAR TO C1R AND C1S.
 CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D16492; BAA03944.1; -.
 DR HSSP: P00736; 1AQP.
 DR MEROPS: S01.198; -.
 DR MGD: MGI:88492; Maspl.
 DR InterPro: IPR000152; -.
 DR InterPro: IPR000436; -.
 DR InterPro: IPR000561; -.
 DR InterPro: IPR000859; -.
 DR InterPro: IPR001254; -.
 DR InterPro: IPR001314; -.
 DR InterPro: IPR001881; -.
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00084; sushi; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 1.
 KW Hydrolyase; Complement pathway; Serine protease; Protease;
 KW Glycoprotein; Sushi; Repeat; Signal; EGF-like domain; Hydroxylation.
 FT SIGNAL 1 24
 FT CHAIN 25 704 COMPLEMENT-ACTIVATING COMPONENT OF
 FT RA-REACTIVE FACTOR (P100).
 FT CHAIN 25 453 70 KDA CHAIN OF P100 (P70).
 FT CHAIN 454 704 29 KDA CHAIN OF P100 (P29).
 FT DOMAIN 25 143 CUB.
 FT DOMAIN 144 187 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 190 302 CUB.
 FT DOMAIN 305 368 SUSHI 1.
 FT DOMAIN 371 438 SUSHI 2.
 FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 557 557 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 651 651 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT MOD_RES 164 164 HYDROXYLATION (POTENTIAL).
 FT DISULFID 78 96 POTENTIAL.
 FT DISULFID 148 162 POTENTIAL.
 FT DISULFID 158 171 POTENTIAL.
 FT DISULFID 173 186 POTENTIAL.
 FT DISULFID 190 217 POTENTIAL.
 FT DISULFID 247 265 POTENTIAL.
 FT DISULFID 306 354 POTENTIAL.
 FT DISULFID 334 367 POTENTIAL.
 FT DISULFID 372 419 POTENTIAL.
 FT DISULFID 402 437 POTENTIAL.
 FT DISULFID 441 577 INTERCHAIN (POTENTIAL).
 FT DISULFID 619 636 POTENTIAL.
 FT DISULFID 647 677 POTENTIAL.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 704 AA; 79895 MW; 71F44F3012D2C67F CRC64;
 Query Match 8.08; Score 147.5; DB 1; Length 704;
 Best Local Similarity 22.58; Pred. No. 0.00013;
 Matches 75; Conservative 59; Mismatches 128; Indels 71; Gaps 18;
 QY 50 VTSGN-----GSIHSPKPPHTYPRNNVLVRLVAVDENVRIOQTDEREGLEDPEDDI 103
 Db 188 VECGSLFTORTGTTSTDPYNPYKSECSY-TIDLEEGFWSIQEDIFDIEDHPEVP 246
 QY 104 KYDFVEVEPPSDGSLVRCWGSGTVFGKTSKGNHIRFVSDVEYFSEPGFCHYHSII 163
 Db 247 CPDYDIKIKAGS--KVGWPFCEKES-PEPISQTSHSVQILFRSDNSGENR-GWRLSY--- 299
 QY 164 MPQVETTSVSLPP-----SSLSLDLLNNAVTAFTLEE-----LIRYLEP 205
 Db 300 --RAAGNECPKLPQPVYGVKIEPSQAVYSFQDLVSCDTGYKVLKDNQVMDTFOIECLKD 357
 QY 206 DRQVVDLSDLVKPTWOLLGKAFLYGKSKVY-----NLNLLKEVKLYSCTPRNFSVSI 259
 Db 358 GAWSNKI-----PTCKIVDCGAPAGLKHGLVTFSTRNLTYYKSEIR-YSCQPYQYK-- 409
 QY 260 REELKRTDTIF-----WFGCLLVKRCGGNCACCLHNCNECQCVPKRVTKYKHYEVLQR 312
 Db 410 ---LHNTGGVYTCSAHGHTWTNKVL-KR---SLPTCLPVCG----VPRFSRQISRIENGR 458
 QY 313 PKTGVKGLHLSLTDVALEHHEECDCVCRNAGG 345
 Db 459 P--AQKG---TMPWIAMLSHLNGQPFCCGSLG 486

Search completed: September 5, 2001, 10:56:05
 Job time: 122 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:03 ; Search time 24.64 Seconds
(without alignments)
1852.485 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 1848
Sequence: 1 MLLGLLLLSALAGORTGT.....DVALEHHECDVCVRGNAGG 345

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_16.*
- 1: sp.archaea.*
 - 2: sp.bacteria.*
 - 3: sp.fungi.*
 - 4: sp.human.*
 - 5: sp.invertebrate.*
 - 6: sp.mammal.*
 - 7: sp.mhc.*
 - 8: sp.organella.*
 - 9: sp.phage.*
 - 10: sp.plant.*
 - 11: sp.rodent.*
 - 12: sp.unclassified.*
 - 13: sp.vertebrate.*
 - 14: sp.virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1848	100.0	345	11 Q9QY71	Q9QY71 mus musculus
2	1819	98.4	345	11 Q9JHV8	Q9JHV8 mus musculus
3	1801	97.5	345	11 Q9EQX6	Q9EQX6 rattus norv
4	1667	90.2	345	4 Q9UL22	Q9UL22 homo sapien
5	1664	90.0	345	4 Q9NRA1	Q9NRA1 homo sapien
6	1552	84.0	345	13 Q9I946	Q9I946 gallus gall
7	754	40.8	370	11 Q9EQT1	Q9EQT1 rattus norv
8	752	40.7	370	4 Q9GZP0	Q9GZP0 homo sapien
9	192.5	10.4	691	13 Q57658	Q57658 gallus gall
10	191	10.3	977	13 Q91925	Q91925 xenopus lae
11	186	10.1	735	13 Q57381	Q57381 xenopus lae
12	185	10.0	926	4 Q9Q000	Q9Q000 homo sapien
13	185	10.0	1015	4 Q9Y6L7	Q9Y6L7 homo sapien
14	183	9.9	1012	11 Q9WVM6	Q9WVM6 mus musculus
15	181	9.8	241	11 Q9Z135	Q9Z135 rattus norv
16	176	9.5	704	4 Q9H2E1	Q9H2E1 homo sapien
17	174.5	9.4	3623	4 Q60494	Q60494 brachydanio
18	174	9.4	1022	13 Q57460	Q57460 gallus gall
19	169	9.1	1008	13 Q9DER7	Q9DER7 gallus gall

20	168	9.1	921	11 Q9QX38	Q9QX38 rattus norv
21	165	8.9	1013	11 Q62381	Q62381 mus musculus
22	165	8.9	1019	13 Q57382	Q57382 xenopus lae
23	162	8.8	1013	4 Q43897	Q43897 homo sapien
24	162	8.8	1013	4 Q9NQS4	Q9NQS4 homo sapien
25	160	8.7	3623	11 Q70244	Q70244 rattus norv
26	158	8.5	3620	6 Q9TU53	Q9TU53 canis famil
27	157	8.5	415	4 Q9UKZ9	Q9UKZ9 homo sapien
28	155.5	8.4	555	4 Q9H2E2	Q9H2E2 homo sapien
29	155.5	8.4	901	4 Q9H2E4	Q9H2E4 homo sapien
30	155.5	8.4	901	4 Q9H2D5	Q9H2D5 homo sapien
31	155.5	8.4	906	4 Q9H2E3	Q9H2E3 homo sapien
32	155.5	8.4	906	4 Q9H2D4	Q9H2D4 homo sapien
33	153.5	8.3	326	11 Q35251	Q35251 rattus norv
34	150.5	8.1	686	13 Q9DGC2	Q9DGC2 cyprinus ca
35	149	8.1	212	11 Q09020	Q09020 rattus norv
36	149	8.1	701	11 Q9JJS9	Q9JJS9 rattus norv
37	148	8.0	746	5 Q01654	Q01654 halocynthia
38	147.5	8.0	1524	13 Q91674	Q91674 xenopus lae
39	145.5	7.9	686	4 Q9Y270	Q9Y270 homo sapien
40	144	7.8	699	4 Q95570	Q95570 homo sapien
41	144	7.8	699	4 Q9UF09	Q9UF09 homo sapien
42	143.5	7.8	597	5 Q26051	Q26051 paracentrot
43	142	7.7	685	13 Q9DGC1	Q9DGC1 cyprinus ca
44	142	7.7	685	13 Q9DGC0	Q9DGC0 cyprinus ca
45	141.5	7.7	719	13 Q9PVI2	Q9PVI2 triakis scy

ALIGNMENTS

RESULT 1
Q9QY71 PRELIMINARY; PRT; 345 AA.
AC Q9QY71;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE FALLOTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Tsai Y.-J., Lee R.K.-K., Chen Y.-H., Lin S.-P., Cheng W.T.-K.;
RT "CDNA cloning of falloletin from mouse ovary.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117608; AAF22516.1;
DR InterPro; IPR000072;
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38741 MW; 3A58A1F701B84EA2 CRC64;

Query Match 100.0%; Score 1848; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 3.5e-159;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLLGLLLLSALAGORTGT	RAESNLSSKQLQSSDKQNGVQDP	HRVVTISGNGSIHS	60
DB	1	MLLGLLLLSALAGORTGT	RAESNLSSKQLQSSDKQNGVQDP	HRVVTISGNGSIHS	60
QY	61	PKPFPYPRNVLWRLVAYDEN	VRIQLTDERFGLDEPDICKYDFVE	VEPSPDGSVL	120
DB	61	PKPFPYPRNVLWRLVAYDEN	VRIQLTDERFGLDEPDICKYDFVE	VEPSPDGSVL	120
QY	121	GRWCGSGTVPGKTSKGNHIRFV	SDYFPSEPGFCIHYSIIMPQVTT	TSFVLPSPS	180
DB	121	GRWCGSGTVPGKTSKGNHIRFV	SDYFPSEPGFCIHYSIIMPQVTT	TSFVLPSPS	180

```
Db 121 GRWCGSTVPKQTSKGNHIRIRFVSDEYFPEPFGFCIHYSIIMPQVTTTSPSVLPSS 180
Qy 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Db 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Qy 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Db 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Qy 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
Db 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345

RESULT 2
Q9JHV8 PRELIMINARY; PRT; 345 AA.
AC Q9JHV8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
GN PDGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS-WEESTER/NIH;
RA Ding H., Wu X., Kim I., Tam P.P.L., Koh G.Y., Nagy A.;
RT "the mouse pdgfc gene: Dynamic expression in embryonic tissues during
RL organogenesis";
DR Mech. Dev. 0:0-0(2000).
DR EMBL; AF286725; AAF91483.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38886 MW; FA1486BED6D362F8 CRC64;

Query Match 98.4%; Score 1819; DB 11; Length 345;
Best Local Similarity 98.8%; Pred. No. 1.5e-156;
Matches 341; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MLLGLLLLTALAGORTGTAEESLSSKLQSSDKQNGVODPRHVVTTISGNGSIHS 60
Db 1 MLLGLLLLTALAGORTGTAEESLSSKLQSSDKQNGVODPRHVVTTISGNGSIHS 60
Qy 61 PKPHTYPRNMVLRVAVDENVRITQTFDERFGLDEPDDICKYDFVEVEEPSDGSYL 120
Db 61 PKPHTYPRNMVLRVAVDENVRITQTFDERFGLDEPDDICKYDFVEVEEPSDGSYL 120
Qy 121 GRWCGSTVPKQTSKGNHIRIRFVSDEYFPEPFGFCIHYSIIMPQVTTTSPSVLPSS 180
Db 121 GRWCGSTVPKQTSKGNHIRIRFVSDEYFPEPFGFCIHYSIIMPQVTTTSPSVLPSS 180
Qy 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Db 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Qy 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Db 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Qy 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
Db 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345

RESULT 3
Q9EQX6 PRELIMINARY; PRT; 345 AA.
AC Q9EQX6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN RSCDGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=KIDNEY;
RA Hamada T., Oi-Rei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDGF-B, a Novel Growth Factor Homologous to
RT SCDGF/PDGF-C/fallotelin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
DR EMBL; AB033830; BAB19969.1; -.
SQ SEQUENCE 345 AA; 38734 MW; F296DA6E9B765D10 CRC64;

Query Match 97.5%; Score 1801; DB 11; Length 345;
Best Local Similarity 96.8%; Pred. No. 6.2e-155;
Matches 334; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MLLGLLLLTALAGORTGTAEESLSSKLQSSDKQNGVODPRHVVTTISGNGSIHS 60
Db 1 MLLGLLLLTALAGORTGTAEESLSSKLQSSDKQNGVODPRHVVTTISGNGSIHS 60
Qy 61 PKPHTYPRNMVLRVAVDENVRITQTFDERFGLDEPDDICKYDFVEVEEPSDGSYL 120
Db 61 PKPHTYPRNMVLRVAVDENVRITQTFDERFGLDEPDDICKYDFVEVEEPSDGSYL 120
Qy 121 GRWCGSTVPKQTSKGNHIRIRFVSDEYFPEPFGFCIHYSIIMPQVTTTSPSVLPSS 180
Db 121 GRWCGSTVPKQTSKGNHIRIRFVSDEYFPEPFGFCIHYSIIMPQVTTTSPSVLPSS 180
Qy 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Db 181 LSLDLLNNAVTAFSTLEELIRYLEPRDQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240
Qy 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Db 241 LKEEVKLYSCTPRNFSVSIRELKRDTDTTFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300
Qy 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345
Db 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345

RESULT 4
Q9UL22 PRELIMINARY; PRT; 345 AA.
AC Q9UL22;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SECRETORY GROWTH FACTOR-LIKE PROTEIN FALLOTEIN (SPINAL CORD-DERIVED
GN GROWTH FACTOR).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Tsai Y.J., Lee R.K.K., Lin S.P.;
RT "Fallotelin, a novel growth factor like gene identified in human
```

uterus.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBDJ databases.
RN [2]

RC SEQUENCE FROM N.A.

RP TISSUE=BRAIN;

RA MEDLINE=20317014; PubMed=10858496;

RX Hamada T., Ui-Tei K., Miyata Y.;

RT "A novel gene derived from developing spinal cords, SCDGF, is a unique

member of the PDGF/VEGF family.";

RL FEBS Lett. 475:97-102(2000).

DR EMBL; AF091434; AAF00049.1; -

DR EMBL; AB033831; BAB03266.1; -

DR InterPro; IPR000072; -

DR InterPro; IPR000859; -

DR Pfam; PF00341; PDGF; 1.

DR Pfam; PF00431; CUB; 1.

DR PROSITE; PS01180; CUB; 1.

DR PROSITE; PS02078; PDGF_2; 1.

DR SMART; SM00042; CUB; 1.

SQ SEQUENCE 345 AA; 39029 MW; CDE9E51F40633E78 CRC64;

Query Match 90.2%; Score 1667; DB 4; Length 345;

Best Local Similarity 87.0%; Pred. No. 8.4e-143;

Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLTALAGORTGTRAESNLSSKQLQSSDKQNGVQDPRHRYVTVTSNGSIHS 60

DB 1 MSLFLLLTALAGORGTQAESNLSSKQFSSNKEQNGVQDPQHERIITVTSNGSIHS 60

QY 61 PKPHTYPRNMVLRVAVDENVRIOITFDERFGLDEPDICKYDFVEVEEPSDGSVL 120

DB 61 PRPHTYPRNTVLVRLVAVENWVIOITFDERFGLDEPDICKYDFVEVEEPSDGTIL 120

QY 121 GRWCGSGVPGKQTSKGNHRIIRFVSDEYFPSEPGFCIHYSIIMPQVTTTSPSVLPSS 180

DB 121 GRWCGSGVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPSVLPSSA 180

QY 181 LSLDLLNNAVTAFSTLEELIRYLEPDRQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240

DB 181 LPDLLNNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKSRVVDLNL 240

QY 241 LKEEVKLYSCTPRNFSVSIRELKRITDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300

DB 241 LTEEVRLYSCTPRNFSVSIRELKRITDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPSK 300

QY 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345

DB 301 VTKKYHEVLQRPKTVGRLGHLSLTDVALEHHEECDCVCRGSTGG 345

RESULT 5

Q9NRA1

ID Q9NRA1 PRELIMINARY; PRT; 345 AA.

AC Q9NRA1;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE PLATELET-DERIVED GROWTH FACTOR C.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RC SEQUENCE FROM N.A.

RP TISSUE=LUNG;

RA Li X., Ponten A., Aase K., Karlsson L., Abramsson A., Uutela M.,

RA Backstrom G., Hellstrom M., Bostrom H., Li H., Soriano P.,

RA Betsholtz C., Heldin C.-H., Alitalo K., Ostman A., Eriksson U.;

RT "PDGF-C is a novel protease-activated ligand for the PDGF alpha

receptor.";

RL Nat. Cell Biol. 0:0-0(2000).

DR EMBL; AF244813; AAF80597.1; -

DR InterPro; IPR000072; -

DR InterPro; IPR000859; -

DR Pfam; PF00341; PDGF; 1.

DR Pfam; PF00431; CUB; 1.

DR PROSITE; PS01180; CUB; 1.

DR PROSITE; PS02078; PDGF_2; 1.

DR SMART; SM00042; CUB; 1.

SQ SEQUENCE 345 AA; 39043 MW; 590889CEA55CC5EA CRC64;

Query Match 90.0%; Score 1664; DB 4; Length 345;

Best Local Similarity 86.7%; Pred. No. 1.6e-142;

Matches 299; Conservative 28; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLLGLLLLTALAGORTGTRAESNLSSKQLQSSDKQNGVQDPRHRYVTVTSNGSIHS 60

DB 1 MSLFLLLTALAGORGTQAESNLSSKQFSSNKEQNGVQDPQHERIITVTSNGSIHS 60

QY 61 PKPHTYPRNMVLRVAVDENVRIOITFDERFGLDEPDICKYDFVEVEEPSDGSVL 120

DB 61 PRPHTYPRNTVLVRLVAVENWVIOITFDERFGLDEPDICKYDFVEVEEPSDGTIL 120

QY 121 GRWCGSGVPGKQTSKGNHRIIRFVSDEYFPSEPGFCIHYSIIMPQVTTTSPSVLPSS 180

DB 121 GRWCGSGVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPSVLPSSA 180

QY 181 LSLDLLNNAVTAFSTLEELIRYLEPDRQVLDLSLYKPTWQLLGKAFYKSKVYNLNL 240

DB 181 LPDLLNNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKSRVVDLNL 240

QY 241 LKEEVKLYSCTPRNFSVSIRELKRITDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPRK 300

DB 241 LTEEVRLYSCTPRNFSVSIRELKRITDTIFWPGCLLVKRCGGNCACCLHNCNCCQVPSK 300

QY 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEECDCVCRGNAGG 345

DB 301 VTKKYHEVLQRPKTVGRLGHLSLTDVALEHHEECDCVCRGSTGG 345

RESULT 6

Q9I946

ID Q9I946 PRELIMINARY; PRT; 345 AA.

AC Q9I946;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE SPINAL CORD-DERIVED GROWTH FACTOR.

GN SCDGF.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OC NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WHITE LECHORN; TISSUE=SPINAL CORD;

RX MEDLINE=20317014; PubMed=10858496;

RA Hamada T., Ui-Tei K., Miyata Y.;

RT "A novel gene derived from developing spinal cords, SCDGF, is a unique

member of the PDGF/VEGF family.";

RL FEBS Lett. 475:97-102(2000).

DR EMBL; AB033829; BAB03265.1; -

DR InterPro; IPR000072; -

DR InterPro; IPR000859; -

DR Pfam; PF00431; CUB; 1.

DR PROSITE; PS01180; CUB; 1.

DR PROSITE; PS02078; PDGF_2; 1.

DR SMART; SM00042; CUB; 1.

SQ SEQUENCE 345 AA; 38940 MW; 97ACEA992BF5128C CRC64;

Query Match

Best Local Similarity 84.0%; Score 1552; DB 13; Length 345;

80.3%; Pred. No. 2.2e-132;

Matches 277; Conservative 37; Mismatches 31; Indels 0; Gaps 0;

```
QY 1 MLLGLLITSAAGORTGTRAESNLSSKQLSSDKEQGVDPDRHVRVVTISGNGSIHS 60
DB 1 MLLGLLITSAAGRRHGAARESDLSKSFSPGAKAQGVDPQHEKIITVTSNGSIHS 60
QY 61 PKFPHYPRNMVLRVAVDENVRIQTLTDFDERFGLDEPDDICKYDFVEVEEPSGVL 120
DB 61 PKFPHYPRNTVLVRLVAVDENVVLIQTLTDFDERFGLDEPDDICKYDFVEVEEPSGTVL 120
QY 121 GRWCGSGTVPGKQTSKGNHIRIRFVSDEYFPPSEPGFCIHYSIMPOVTTTSPSVLPSS 180
DB 121 GRWCGSSVSQKISGNQIRIRFVSDEYFPPSQPGFCIHITLLVPHHTTAPSSSLPPSA 180
QY 181 LSLDLNNAVAFSTLEELIRYLEDRMOVDLSLYKPTWQLLGRAFLYKSKKVVNLNL 240
DB 181 LPLDLNNAVAGFVVEELIRYLEDRWQDLELYRPTWQLLGRAIYHGRKSRVVDLNL 240
QY 241 LKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPRK 300
DB 241 LKEEVRLYSCPTPRNFSVLSREELKRTDTIFWPLCLLVKRCGNCACCHQNCNCCQIPTK 300
QY 301 VTKYHEVLQLRPKTGVKGLHSLTDVALEHHEECDCVCRNAGG 345
DB 301 VTKYHEVLQLRPSGVRGLHSLTDVPLEHHEECDCVCKGNSEG 345
```

RESULT 7

```
Q9EQT1
ID Q9EQT1 PRELIMINARY; PRT; 370 AA.
AC Q9EQT1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE SPINAL-CORD DERIVED GROWTH FACTOR-B.
GN RSCDGF-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamada T., Oi-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDF-B, a Novel Growth Factor Homologous to
RT SCDF/PDGF-C/falotelin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
DR EMBL; AB052170; BAB18920.1; -.
SQ SEQUENCE 370 AA; 42809 MW; 7BE8A251F679BF73 CRC64;
```

Query Match 40.8%; Score 754; DB 11; Length 370;
Best Local Similarity 46.28; Pred. No. 4.1e-60;
Matches 151; Conservative 62; Mismatches 86; Indels 28; Gaps 10;

```
QY 37 EQNGVOD-PRHERVVTISGNGSIHSPKPHYPRNMVLRVAVDENVRIQTLTDFDERG 95
DB 42 ESNHLDLYRRDENIRVTGTCHVQSPRPNSYPRNLLTWRLHS-QEXTRIQIAFDHQFG 100
QY 96 LEDPEDDICKYDFVEVEEPSGDS--VLGRWCGSGTVPGKQTSKGNHIRIRFVSDEYFPPSE 153
DB 101 LEEAENDICRYDFVEVEDVSSTVVRGCGHKEIPRITSRTNQIKITFQSDDYFVAK 160
QY 154 PGFCIHYSII---MPO-----VTET-----TSPSVLPSSLSLDLNNVAFST 195
DB 161 PGFKIYISFVEDFQPEASEINWESVTSFGSVHSPVM-DSTLTADALDKAIAEFT 219
QY 196 LEELIRYLEPDRWQVDSLKYKPTWQLLGRAFLYKSKKVVNLNLLKEEVKLYSCTPRNF 255
DB 220 VEDLLKYFNPAWQDLENLYMDTPRYGRSY-HERKSK-VDLDRLNDVKKRYSCTPRNH 277
QY 256 SVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPRKVTYKHYEVLQLRP-- 313
DB 278 SVNIREELKLTNAVFFPRCLLVQRCGCGCGTLLNWKSCTSCKSSGKTVKHYEVLKFEPEGH 337
```

RESULT 9

```
O57658
ID O57658 PRELIMINARY; PRT; 691 AA.
AC O57658;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
```

```
QY 314 -KTGVKGLHSLTDVALEHHEECDCVC 339
DB 338 FKRRGAKNALVDIQLDHERCDCIC 364
```

RESULT 8

```
Q9GZP0
ID Q9GZP0 PRELIMINARY; PRT; 370 AA.
AC Q9GZP0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE SPINAL-CORD-DERIVED GROWTH FACTOR-B (MSTP036).
GN RSCDGF-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamada T., Oi-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDF-B, a Novel Growth Factor Homologous to
RT SCDF/PDGF-C/falotelin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-AORTA;
RA Liu B., Liu Y.-Q., Wang X.-Y., Zhao B., Sheng H., Zhao X.-W., Liu S.,
RA Xu Y.-Y., Ye J., Song L., Gao Y., Zhang C.-L., Zhang J., Wei Y.-J.,
RA Cao H.-Q., Zhao Y., Liu L.-S., Ding J.-F., Gao R.-L., Wu Q.-Y., Qiang B.-Q.,
RA Yuan J.-G., Liew C.-C., Zhao M.-S., Hui R.-T.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB033832; BAB18903.1; -.
DR EMBL; AF113216; AAG39287.1; -.
SQ SEQUENCE 370 AA; 42848 MW; D387F485E7BB7674 CRC64;
```

Query Match 40.7%; Score 752; DB 4; Length 370;
Best Local Similarity 45.3%; Pred. No. 6.3e-60;
Matches 148; Conservative 59; Mismatches 92; Indels 28; Gaps 9;

```
QY 37 EQNGVOD-PRHERVVTISGNGSIHSPKPHYPRNMVLRVAVDENVRIQTLTDFDERG 95
DB 42 ESNHLDLYRRDETQVKGNGYVQSPRPNSYPRNLLTWRLHS-QENTRIQLVFNQFG 100
QY 96 LEDPEDDICKYDFVEVEEPSGDSVL--GRWCGSGTVPGKQTSKGNHIRIRFVSDEYFPPSE 153
DB 101 LEEAENDICRYDFVEVEDISSETIIRGRWCGHKEVPPRIKSRTNQIKITFKSDDYFVAK 160
QY 154 PGFCIHYSII---MPQVTET-----SPSVLPSSLSLDLNNVAFST 195
DB 161 PGFKIYISLEDFQPAASEINWESVTSFGSVSPSVTDP-TLIADALDKKIAEFT 219
QY 196 LEELIRYLEPDRWQVDSLKYKPTWQLLGRAFLYKSKKVVNLNLLKEEVKLYSCTPRNF 255
DB 220 VEDLLKYFNPEWQEDLENLYMDTPRYGRSY-HDRKSK-VDLDRLNDDAKRYSCTPRNY 277
QY 256 SVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPRKVTYKHYEVLQLRP-- 313
DB 278 SVNIREELKANVFFPRCLLVQRCGCGCGTVNNRSCCTNSGKTVKHYEVLQFEPEGH 337
QY 314 -KTGVKGLHSLTDVALEHHEECDCVC 339
DB 338 IKRRGAKTMALVDIQLDHERCDCIC 364
```


[illegible]

RESULT	11	
ID	057381	PRELIMINARY; PRT; 735 AA.
AC	Q57381;	
DT	01-JUN-1998	(TrEMBLrel. 06, Created)
DT	01-JUN-1998	(TrEMBLrel. 06, Last sequence update)
DT	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)
DE	BONE MORPHOGENETIC PROTEIN 1B.	
GN	BMP-1.	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=OVARY;	
RA	Goodman S., Albano R., Matthews G., Tannahill D., Dale L.;	
RL	Dev. Biol. 0:0-0(0).	
DR	EMBL; Y09660; CAA70853.1; -.	
DR	HSSP; P00736; IAPQ.	
DR	InterPro; IPR000130; -.	
DR	InterPro; IPR000152; -.	
DR	InterPro; IPR000561; -.	
DR	InterPro; IPR000859; -.	
DR	InterPro; IPR001506; -.	
DR	InterPro; IPR001881; -.	
DR	Pfam; PF000008; EGF; 1.	
DR	Pfam; PF00431; CUB; 3.	
DR	Pfam; PF01400; Astacin; 1.	
DR	PRINTS; PR00480; ASTACIN.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.	
DR	PROSITE; PS01180; CUB; 3.	
DR	PROSITE; PS01186; EGF; 2; 1.	
DR	PROSITE; PS01187; EGF-CA; 1.	
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.	
DR	SMART; SM00042; CUB; 1.	

RESULT	11	
ID	057381	PRELIMINARY; PRT; 735 AA.
AC	Q57381;	
DT	01-JUN-1998	(TrEMBLrel. 06, Created)
DT	01-JUN-1998	(TrEMBLrel. 06, Last sequence update)
DT	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)
DE	BONE MORPHOGENETIC PROTEIN 1B.	
GN	BMP-1.	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=OVARY;	
RA	Goodman S., Albano R., Matthews G., Tannahill D., Dale L.;	
RL	Dev. Biol. 0:0-0(0).	
DR	EMBL; Y09660; CAA70853.1; -.	
DR	HSSP; P00736; IAPQ.	
DR	InterPro; IPR000130; -.	
DR	InterPro; IPR000152; -.	
DR	InterPro; IPR000561; -.	
DR	InterPro; IPR000859; -.	
DR	InterPro; IPR001506; -.	
DR	InterPro; IPR001881; -.	
DR	Pfam; PF000008; EGF; 1.	
DR	Pfam; PF00431; CUB; 3.	
DR	Pfam; PF01400; Astacin; 1.	
DR	PRINTS; PR00480; ASTACIN.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.	
DR	PROSITE; PS01180; CUB; 3.	
DR	PROSITE; PS01186; EGF; 2; 1.	
DR	PROSITE; PS01187; EGF-CA; 1.	
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.	
DR	SMART; SM00042; CUB; 1.	

[illegible]

GN	Mus musculus (Mouse)
ON	Eukaryota; Metazoa; Chordata;
OC	Mammalia; Eutheria; Rodentia;
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:38 ; Search time 20.92 Seconds
(without alignments)
999,774 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 345
Sequence: 1 MLLGLLLLSALAGQRTGT.....DVALEHHECDVCVRGNAGG 345

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	100.0	345	21 AAB48658	Mouse zveg3, SEQ
2	345	100.0	345	21 AAY96861	Murine vascular en
3	345	100.0	345	21 AAY84559	A murine platelet-
4	94	27.2	180	21 AAB01427	Murine TANGO 128.
5	52	15.1	113	21 AAB10631	Human VEGF-X prote
6	52	15.1	113	21 AAB10632	Human VEGF-X prote
7	52	15.1	149	21 AAB10642	Human VEGF-X PDGF-
8	52	15.1	227	21 AAB10637	Human VEGF-X prote
9	52	15.1	227	21 AAB10638	Human VEGF-X prote
10	52	15.1	318	21 AAY84558	A fragment of plat
11	52	15.1	339	21 AAB58438	Lung cancer associ

12	52	15.1	345	20 AAY33679	Human VEGF-E prote
13	52	15.1	345	20 AAY41766	Human PRO200 prote
14	52	15.1	345	20 AAY30023	Human vascular end
15	52	15.1	345	21 AAB48657	Human zveg3, SEQ
16	52	15.1	345	21 AAB24250	Human platelet-der
17	52	15.1	345	21 AAB44322	Human PRO200 (UN01
18	52	15.1	345	21 AAB10633	Human RACE generat
19	52	15.1	345	21 AAB10634	Human VEGF-X homol
20	52	15.1	345	21 AAB10635	Human VEGF-X prote
21	52	15.1	345	21 AAB10636	Human VEGF-X prote
22	52	15.1	345	21 AAB10644	Human VEGF-X prote
23	52	15.1	345	21 AAB10650	Human 990126vegX p
24	52	15.1	345	21 AAB10651	Human VEGF-X prote
25	52	15.1	345	21 AAB19578	Human PRO200 (vasc
26	52	15.1	345	21 AAB33414	Human PRO200 prote
27	52	15.1	345	21 AAB24412	Human PRO713 prote
28	52	15.1	345	21 AAB01419	Human TANGO 128.
29	52	15.1	345	21 AAB03003	Human growth facto
30	52	15.1	345	21 AAY96858	Human growth facto
31	52	15.1	345	21 AAY84557	Amino acid sequenc
32	52	15.1	345	21 AAY59285	Bone morphogenic p
33	52	15.1	345	22 AAB50980	Human PRO200 prote
34	52	15.1	345	22 AAB49895	Human PRO200 prote
35	52	15.1	345	22 AAB53074	Human angiogenesis
36	52	15.1	354	21 AAB10640	Human VEGF-X prote
37	52	15.1	374	21 AAB10639	Human VEGF-X prote
38	52	15.1	374	21 AAB10639	Human VEGF-X prote
39	32	9.3	167	21 AAB10652	Human VEGF-X prote
40	32	9.3	168	21 AAB10643	Human VEGF-X CUB-1
41	32	9.3	282	21 AAB10653	Human VEGF-X prote
42	32	9.3	297	21 AAY59286	Bone morphogenic p
43	9	2.6	22	21 AAY84560	Antigenic peptide
44	9	2.6	494	15 AAR62825	Human steroid-21-h
45	8	2.3	49	20 AAY25293	HCV NS5B carboxy-t

ALIGNMENTS

RESULT 1
AAB48658
ID AAB48658 standard; Protein; 345 AA.
AC AAB48658;
XX
DT 09-MAR-2001 (first entry)
XX
DE Mouse zveg3, SEQ ID NO:35.
XX
KW Mouse; zveg3; zveg3 fusion; growth factor homologue; VEGF/PDGF family;
KW murine; CUB domain; PDGF-like activity; mitogenic; osteogenic;
KW neovascularisation; tissue repair; proliferation; differentiation;
KW liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
KW periodontal disease; bone fracture; wound healing; vulnery; ischaemia;
KW immunomodulation; hepatic.
XX
OS Mus musculus.
XX
PN WO200066736-A1.
XX
PD 09-NOV-2000.
XX
PF 03-MAY-2000; 2000WO-US40047.
XX
PR 03-MAY-1999; 99US-0304216.
PR 10-NOV-1999; 99US-0164463.
PR 04-FEB-2000; 2000US-0180169.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;
XX
DR WPI; 2000-687541/67.

DR N-PSDB; AAC81583.
XX Growth factor homologs and the nucleic acids that encode them, useful
PT e.g. for treating liver damage, ischemia, multiple sclerosis and
PT Alzheimer's disease -
XX
PS Disclosure; Page 130-131; 143pp; English.
XX
CC The invention relates to the human growth factor homologue zvegfg4
CC (AAB48653), and nucleic acids encoding it (AAC81555). Zvegfg4 is a member
CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
CC growth factor) family. Zvegfg4 has a growth factor domain (AAB48654)
CC characterised by a PDGF cysteine knot structure, and a CUB domain
CC (AAB48655) which has a beta barrel structure. Zvegfg4 has PDGF-like
CC activity, having mitogenic activity on fibroblasts, vascular smooth
CC muscle cells and pericytes, and has also been shown to stimulate bone
CC growth. The invention also relates to fusion proteins comprising human
CC zvegfg4 or fragments thereof, particularly human zvegfg4/human zvegfg3
CC fusions; expression constructs and host cells comprising human zvegfg4
CC nucleic acids; the recombinant expression of human zvegfg4; an antibody
CC which binds to human zvegfg4 or a fragment thereof; a method of activating
CC a cell-surface PDGF receptor using a zvegfg4-derived polypeptide; a
CC method of modulating the proliferation, differentiation, migration or
CC metabolism of bone cells, comprising exposing bone cells to
CC zvegfg4-derived polypeptides; and a method of detecting a genetic
CC abnormality in the zvegfg4 gene of a patient. Zvegfg4 proteins and derived
CC fragments may be used to stimulate tissue development or repair, or
CC cellular differentiation or proliferation. They are particularly used for
CC the treatment or repair of liver damage, and may also be used to
CC modulate neurite growth (e.g., in the treatment of Alzheimer's disease or
CC multiple sclerosis). Due to their osteogenic activity, they may be used
CC in the treatment of periodontal disease and fractures. They may also be
CC used to enhance expansion and mobilisation of haematopoietic stem cells
CC and endothelial precursor stem cells, which may be useful in the
CC treatment of ischaemia, in wound healing, and in the modulation of the
CC immune system. The present sequence represents mouse zvegfg3.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLGLLLTTSALAGQRTGTRAESNLSSKLQSSDKKEQNGVQDPRHVRVVTISGNSIHS 60
DB 1 mllglllttsalagrtgtraesnlssklqssdkkeqngvqdprrhvrvtisgnsihs 60
QY 61 PKFPHYPRNMVLVRLVAVDENVRILQTFDERFGLDEPDDICKYDFVEVEPSPDGSVL 120
DB 61 pkfphypnmvlvrlvavdenvrilqtfderfgldedddickdyfvevepsdgsvl 120
QY 121 GRWCSSGVPGKQTSKGNHIRFVSDFFPPEPGCIHYSIIMPOVTTSPSVLPSS 180
DB 121 grwcsgvpgkqtskgnhირფვსდფფპეპგციჰყსიიმპოვტტსპსლპსს 180
QY 181 LSLDLNNAVTAFTLEELIRYLEPDRWQVDLSLKYPTWLLGKFLGKSKVYNLNL 240
DB 181 lsldlnnavtaftleelirylepdrwqvdlslykptwllgkflgkkskvvynlnl 240
QY 241 LKEEVKLYSCTPRNFSVSIREELKRTDTTFWPGCLLVKRCGGNCACCLHNCBCQVPRK 300
DB 241 lkeevklyscprnfsvsireelkrtdttfwpgcllvkrcggncacclhncbcqvprk 300
QY 301 VTKKYHEVLQRPKTVGKGLHLSLTDVALEHHEBCDCVCRGNAGG 345
DB 301 vtkeyhevlqrpktvgkglhlsldvalehhebcdcvcrgnagg 345

RESULT 2

AAY96861

ID AAY96861 standard; Protein; 345 AA.

XX

AC AAY96861;
XX 26-SEP-2000 (first entry)
DT Murine vascular endothelial growth factor homologue, ZVEGF3.
DE
XX
XX Vascular endothelial growth factor; homologue; zvegfg3; CUB domain;
KW Cysteine knot; platelet-derived growth factor; PDGF; neuropilin;
KW chromosome 4q28.3; cytostatic; anti-psoriatic; anti-inflammatory;
KW anti-diabetic; ophthalmological; anti-rheumatic; anti-arthritis;
KW vulnery.
XX
XX Mus musculus.
OS
XX WO200034474-A2.
PN
XX
XX 15-JUN-2000.
PD
XX
XX 07-DEC-1999; 99WO-US28968.
PF
XX
XX 07-DEC-1998; 98US-0207120.
PR
XX 06-JUL-1999; 99US-0142576.
PR
XX 21-OCT-1999; 99US-0161653.
PR
XX 12-NOV-1999; 99US-0165255.
PR
XX (ZYMO) ZYMOGENETICS INC.
PA
XX Gao Z, Hart CE, Piddington CS, Sheppard PO, Shoemaker KE;
PI Gilbertson DG, West JW;
PI
XX WPI; 2000-4234420/36.
DR N-PSDB; AAA51527.
DR
XX Novel zvegfg3 polypeptides and nucleotides encoding them useful for
PT stimulating growth of smooth muscle cells and fibroblasts comprising an
PT epitope bearing portion of a specific amino acid sequence
XX
XX Claim 1; Page 169-170; 173pp; English.
XX
XX This shows a murine ZVEGF3 a novel vascular endothelial growth factor
CC homologue. Polypeptides comprising an epitope-bearing portion human or
CC murine ZVEGF3 are claimed. The growth factors comprise a growth factor
CC domain and a CUB domain (generic sequence motifs are shown in AAY96859
CC and AAY96860). The growth factor domain is characterized by an
CC arrangement of cysteine residues and beta-strands that is characteristic
CC of the "cysteine knot" structure of the platelet-derived growth factor
CC (PDGF) family. The CUB domain shows homology to CUB domains in
CC neuropilins, human bone morphogenetic protein-1, porcine seminal plasma
CC protein, bovine acidic seminal fluid protein and Xenopus laevis
CC toll-like protein. Structural analysis and homology predict that
CC ZVEGF3 polypeptides complex with a second polypeptide to form multimeric
CC proteins. The human zvegfg3 gene has been mapped to chromosome 4q28.3.
CC ZVEGF3 is useful for stimulating the growth of fibroblasts or smooth
CC muscle cells, for activating cell surface PDGF-alpha receptor and for
CC inhibiting PDGF-alpha receptor mediated cellular processes. ZVEGF3 is
CC useful for regulating (post-development) organ growth, regeneration and
CC maintenance, as well as tissue maintenance and repair processes. ZVEGF3
CC antagonists are useful for treating cancer, rheumatoid arthritis,
CC diabetic retinopathy, ischemic limb disease, peripheral vascular
CC disease, myocardial ischemia, vascular intimal hyperplasia,
CC atherosclerosis, wound healing, chronic liver disease and haemangioma
CC formation. ZVEGF3 can also be used to modulate neurite growth and
CC development of the nervous system, and for treating neurodegenerative
CC diseases.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLGLLLTTSALAGQRTGTRAESNLSSKLQSSDKKEQNGVQDPRHVRVVTISGNSIHS 60

Db 1 mlllglllltsalagrtgtraesnlssklssdkeqngvqdrhervvtisngsihs 60
 QY 61 PKPHTYPRNMVLRVAVDENVRVQLTDFERFGLUEDPEDDICKYDFVEVEEPSDGSVL 120
 Db 61 pkfphypnrmvrlvavdenvrqltderfgleddedickdyfveveepsdgsvl 120
 QY 121 GRWCGSTVPKQTSKGNHIRFVSDEYFPSPGFCIHYSIIMPQVETSPSVLPSS 180
 Db 121 grwcgstvpkgtskgnhirlfvsdeyfpsepgfcihysilmpqvettspsvlpss 180
 QY 181 LSLDLLNNAVTAFTLEELIRYLEPRQWQVLDLSLKYPTWQLLGKAFLYGKSKVNNL 240
 Db 181 lsldllnnnavtafstleelirylepdrwqvdlslykptwqlgkaflygkkskvnlnl 240
 QY 241 LKEEVKLYSCTPRNFVSIRREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPRK 300
 Db 241 lkeevklyscprnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqcvprk 300
 QY 301 VTKKYHEVLQRPKTGVKGLHKSITDVALEHHEECDCVCRGNAGG 345
 Db 301 vtkeyhevlqlrpkgtgkglhksitdvalehheecdcvcrgnagg 345

RESULT 3

AY84559
 ID AAY84559 standard; Protein; 345 AA.
 AC AAY84559;
 XX
 XX 25-JUL-2000 (first entry)
 XX
 DE A murine platelet-derived growth factor C (PDGF-C).
 XX platelet-derived growth factor C; PDGF-C; cell proliferation;
 KW growth factor; heparin; connective tissue; wound healing; VEGF-F;
 KW fibroblast mitogenesis; PDGF alpha receptor activation; tumour growth;
 KW choriocarcinoma; Wilms tumour; megakaryoblastic leukaemia;
 KW lung carcinoma; erythroleukemia; tissue remodelling.
 XX
 OS Mus sp.
 XX
 XX WO200018212-A2.
 PN
 XX
 PD 06-APR-2000.
 XX
 XX 30-SEP-1999; 99WO-0522668.
 XX
 PR 30-SEP-1998; 98US-0102461.
 PR 12-NOV-1998; 98US-0108109.
 PR 03-DEC-1998; 98US-0110749.
 PR 18-DEC-1998; 98US-0113002.
 PR 21-MAY-1999; 99US-0135426.
 PR 15-JUL-1999; 99US-0144022.
 XX

(LUDW-) LUDWIG INST CANCER RES.
 (UYHE-) UNIV HELSINKI LICENSING LTD.

PI Eriksson U, Aase K, Lee X, Ponten A, Uutela M, Allitalo K;
 PI Oestman A, Heldin C, Betsholtz C;

XX WPI; 2000-292954/25.
 DR N-PSDB; AAA12525.

XX Novel DNA encoding PDGF-C useful to stimulate or enhance proliferation,
 PT differentiation, growth and motility of cells expressing the PDGF-C
 PT receptor

XX Claim 27; Fig 6; 135pp; English.

PS The present sequence represents murine platelet-derived growth factor C
 XX (PDGF-C) (formally designated VEGF-F). PDGF-C polypeptides have the
 CC ability to stimulate and enhance proliferation or differentiation,
 CC

CC and/or growth or motility of cells expressing a PDGF-C receptor.
 CC PDGF-C polypeptides can be used in pharmaceuticals for promoting cell
 CC proliferation, preferably in combination with one other growth factor
 CC and heparin. Pharmaceuticals comprising PDGF-C polypeptides can also
 CC be used for stimulating connective tissue or wound healing. The
 CC PDGF-C polypeptide can be enzymatically processed to generate the active
 CC truncated form of PDGF-C and used to regulate the receptor-binding
 CC specificity of PDGF-C. PDGF-C can also be used to promote fibroblast
 CC mitogenesis in a mammal and to induce PDGF alpha receptor activation.
 CC PDGF-C antagonists can be used to inhibit tumour growth of a tumour
 CC expressing PDGF-C in a mammal. Specific types of human tumours, e.g.
 CC choriocarcinoma, Wilms tumour, megakaryoblastic leukaemia, lung carcinoma
 CC and erythroleukemia, can be identified by testing for expression of
 CC PDGF-C. PDGF-C antagonists can also be used to inhibit tissue
 CC remodelling during invasion of tumour cells into a normal population of
 CC cells. Antagonists can also be used to treat fibrotic conditions,
 CC especially found in the lung, kidney or liver.

XX Sequence 345 AA;

Query Match 100.0%; Score 345; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLLLTSALAGQRTGTRAESNLSSKLSSDKEQNGVQDRHERVVTISNGSIHS 60
 Db 1 mlllglllltsalagrtgtraesnlssklssdkeqngvqdrhervvtisngsihs 60
 QY 61 PKPHTYPRNMVLRVAVDENVRVQLTDFERFGLUEDPEDDICKYDFVEVEEPSDGSVL 120
 Db 61 pkfphypnrmvrlvavdenvrqltderfgleddedickdyfveveepsdgsvl 120
 QY 121 GRWCGSTVPKQTSKGNHIRFVSDEYFPSPGFCIHYSIIMPQVETSPSVLPSS 180
 Db 121 grwcgstvpkgtskgnhirlfvsdeyfpsepgfcihysilmpqvettspsvlpss 180
 QY 181 LSLDLLNNAVTAFTLEELIRYLEPRQWQVLDLSLKYPTWQLLGKAFLYGKSKVNNL 240
 Db 181 lsldllnnnavtafstleelirylepdrwqvdlslykptwqlgkaflygkkskvnlnl 240
 QY 241 LKEEVKLYSCTPRNFVSIRREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPRK 300
 Db 241 lkeevklyscprnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqcvprk 300
 QY 301 VTKKYHEVLQRPKTGVKGLHKSITDVALEHHEECDCVCRGNAGG 345
 Db 301 vtkeyhevlqlrpkgtgkglhksitdvalehheecdcvcrgnagg 345

RESULT 4

AAB01427
 ID AAB01427 standard; Protein; 180 AA.

XX AAB01427;

XX 20-OCT-2000 (first entry)

XX Murine TANGO 128.

XX TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
 KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
 KW inflammatory bowel disease; septic shock; ulcerative colitis;
 KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
 KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
 KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
 KW systemic lupus erythematosus; transgenic animal; diagnosis;
 KW prognosis; prophylactic; therapeutic; mouse.

OS Mus musculus.

XX Key Location/Qualifiers
 FT Misc-difference 129

FT Misc-difference 167 /note= "Unidentified amino acid"
 FT /note= "Unidentified amino acid"
 FT Misc-difference 172
 FT /note= "Unidentified amino acid"
 XX

PN WO200039284-A1.

XX 06-JUL-2000.

XX 23-DEC-1999; 99WO-US31025.

XX 30-DEC-1998; 98US-0223546.

XX (MILL-) MILLENNIUM PHARM INC.

XX Holtzman DA;

XX WPI; 2000-465743/40.

DR N-PSDB; AAA47478.

XX Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
 PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
 PT arthritis, psoriasis and autoimmune diseases

XX Claim 8; Fig 26; 209pp; English.

XX Nucleic acids encoding TANGO polypeptides are useful as modulating
 CC agents for regulating cellular processes like asthma, graft
 CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
 CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
 CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
 CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
 CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
 CC lupus erythematosus. The nucleic acids are also useful for producing
 CC transgenic animals and the TANGO polypeptides themselves. Partial
 CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
 CC forensic biology, for diagnostic assays, prognostic assays,
 CC pharmacogenomics and for monitoring clinical trials. TANGO
 CC polypeptides are suitable for both prophylactic and therapeutic
 CC methods for treating a subject at risk of a disorder or having a
 CC disorder associated with aberrant TANGO expression. A wide range
 CC of cellular disorders can be treated.

XX Sequence 180 AA;

Query Match 27.2%; Score 94; DB 21; Length 180;

Best Local Similarity 100.0%; Pred. No. 5.9e-82;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 QRTGTRAESNLSSKQLQLSSDKQNGVQDPRHRRVVTISGNGSTHSPKFPHTYPRNMVLVW 75

Db 14 qrtgtraesnlsskqlqlssdkqngvqdprrhervvtisngsthspskfphtyprnmvlw 73

Qy 76 RLVAVDENVRIQLTFDERFGLEDPEDDICKYDFV 109

Db 74 rlvaundenrqlttderfgledpeddickdyfv 107

RESULT 5

AAB10631

ID AAB10631 standard; Protein; 113 AA.

XX AAB10631;

XX 19-JAN-2001 (first entry)

XX Human VEGF-X protein fragment #1.

XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.

XX Homo sapiens.

XX WO200037641-A2.

XX 29-JUN-2000.

XX 21-DEC-1999; 99WO-US30503.

XX 23-DEC-1998; 98GB-0028377.

XX 18-MAR-1999; 99US-0124967.

XX 08-NOV-1999; 99US-0164131.

XX (JANC) JANSSEN PHARM NV.

XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJJ, Gosiewska A;

PI Dhanaraj SN, Xu J;

XX WPI; 2000-442669/38.

DR N-PSDB; AAA71936.

XX New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -

XX Disclosure; Fig 1; 127pp; English.

XX This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGF-X protein
 CC described in the method of the invention.

XX Sequence 113 AA;

Query Match 15.1%; Score 52; DB 21; Length 113;

Best Local Similarity 100.0%; Pred. No. 5.9e-42;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 LYSECTPRNFVSIREELKRTDTIFWPGCLLVKRGNGCACCLHNCNQCQVP 298

Db 15 lysctprnfsvsireelkrttdtlfpgcllvkrgngcaccclhncnecqcvp 66

RESULT 6

AAB10632

ID AAB10632 standard; Protein; 113 AA.

XX AAB10632;

XX 19-JAN-2001 (first entry)

XX Human VEGF-X protein fragment #2.

XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.
 XX
 OS Homo sapiens.
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 XX
 PR 18-MAR-1999; 99US-0124967.
 PR
 PR 08-NOV-1999; 99US-0164131.
 XX
 XX (JANC) JANSSEN PHARM NV.
 PA
 XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 PI
 XX
 DR WPI: 2000-442669/38.
 DR N-PSDB; AAA71986.
 XX
 XX New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 PT
 XX Disclosure; Fig 2; 127pp; English.
 PS
 XX This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents the human VEGF-X protein
 CC described in the method of the invention.
 XX
 SQ Sequence 113 AA;

Query Match 15.1%; Score 52; DB 21; Length 113;
 Best Local Similarity 100.0%; Pred. No. 5.9e-42;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 247 LYSCTPRNFVSIRBELKRTDTIFWPGCLLVKRCGGNCACCLHNCQCQVP 298
 |||||
 Db 15 lysctprnfsvsireelkrttdtiffpgcllvkrcggncacclhncqcqvp 66
 |||||
 RESULT 7
 AAB10642
 ID AAB10642 standard; Protein; 149 AA.
 XX
 AC AAB10642;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human VEGF-X PDGF-like domain protein.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.

XX Homo sapiens.
 OS
 XX
 PN WO200037641-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US30503.
 XX
 PR 22-DEC-1998; 98GB-0028377.
 PR
 PR 18-MAR-1999; 99US-0124967.
 PR
 PR 08-NOV-1999; 99US-0164131.
 XX
 XX (JANC) JANSSEN PHARM NV.
 PA
 XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
 PI Dhanaraj SN, Xu J;
 PI
 XX
 DR WPI: 2000-442669/38.
 DR N-PSDB; AAA71986.
 XX
 XX New vascular endothelial growth factor protein, useful for treating or
 PT preventing diseases associated with inappropriate angiogenesis activity
 PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
 PT
 XX Disclosure; Fig 24; 127pp; English.
 PS
 XX This invention describes a novel vascular endothelial growth factor-X
 CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
 CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
 CC antidiabetic activity and acts as an angiogenesis and vascularization
 CC regulator. An antisense molecule of the invention is useful for treating
 CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
 CC retinopathy by inhibiting angiogenic activity or inappropriate
 CC vascularization including formation and proliferation of new blood
 CC vessels, growth and development of tissues, tissue regeneration and organ
 CC and tissue repair in a subject. The products of the invention are useful
 CC for preparing medicaments for treating wounds such as dermal ulcers,
 CC pressure sores, venous sores, diabetic ulcers and burns and to promote
 CC skin graft growth, tissue repair, proliferation of new blood vessels,
 CC tissue regeneration and organ repair by promoting angiogenic activity or
 CC vascularization. This sequence represents a human VEGF-X protein
 CC PDGF-like domain which can be expressed in E. coli systems and which is
 CC described in the method of the invention.
 XX
 SQ Sequence 149 AA;

Query Match 15.1%; Score 52; DB 21; Length 149;
 Best Local Similarity 100.0%; Pred. No. 7.4e-42;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 247 LYSCTPRNFVSIRBELKRTDTIFWPGCLLVKRCGGNCACCLHNCQCQVP 298
 |||||
 Db 51 lysctprnfsvsireelkrttdtiffpgcllvkrcggncacclhncqcqvp 102
 |||||
 RESULT 8
 AAB10637
 ID AAB10637 standard; Protein; 227 AA.
 XX
 AC AAB10637;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Human VEGF-X protein #2.
 XX
 KW VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
 KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
 KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
 KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
 KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
 KW venous sore; diabetic ulcer; burns; skin graft growth.

```

XX OS Homo sapiens.
XX PN WO200037641-A2.
XX PD 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-US30503.
XX PR 22-DEC-1998; 98GB-0028377.
XX PR 18-MAR-1999; 99US-0124967.
XX PR 08-NOV-1999; 99US-0164131.
XX PA (JANC ) JANSSEN PHARM NV.
XX XX
XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX DR N-PSDB; AAA71981.
XX DR WPI; 2000-442669/38.
XX XX
XX PT New vascular endothelial growth factor protein, useful for treating or
XX PT preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS Disclosure; Fig 17; 127pp; English.
XX XX
XX CC This invention describes a novel vascular endothelial growth factor-X
XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX CC antidiabetic activity and acts as an angiogenesis and vascularization
XX CC regulator. An antisense molecule of the invention is useful for treating
XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX CC retinopathy by inhibiting angiogenic activity or inappropriate
XX CC vascularization including formation and proliferation of new blood
XX CC vessels, growth and development of tissues, tissue regeneration and organ
XX CC and tissue repair in a subject. The products of the invention are useful
XX CC for preparing medicaments for treating wounds such as dermal ulcers,
XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote
XX CC skin graft growth, tissue repair, proliferation of new blood vessels,
XX CC tissue regeneration and organ repair by promoting angiogenic activity or
XX CC vascularization. This sequence represents a human VEGF-X protein
XX CC described in the method of the invention.
XX SQ Sequence 227 AA;

Query Match 15.1%; Score 52; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 LYSCTPRNFVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCQCV 298
Db 15 lysctprnfsvsireelkrttdtiffpgcllvkrcgncacclhncncqcvp 66

RESULT 9
AAB10638
ID AAB10638 standard; Protein; 227 AA.
AC AAB10638;
XX XX
XX DT 19-JAN-2001 (first entry)
XX DE Human VEGF-X protein #3.
XX XX
XX KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
XX KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
XX KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
XX KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
XX KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
XX KW venous sore; diabetic ulcer; burns; skin graft growth.
XX XX

```

```

OS Homo sapiens.
XX PN WO200037641-A2.
XX PD 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-US30503.
XX PR 22-DEC-1998; 98GB-0028377.
XX PR 18-MAR-1999; 99US-0124967.
XX PR 08-NOV-1999; 99US-0164131.
XX PA (JANC ) JANSSEN PHARM NV.
XX XX
XX PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX DR N-PSDB; AAA71982.
XX DR WPI; 2000-442669/38.
XX XX
XX PT New vascular endothelial growth factor protein, useful for treating or
XX PT preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX PS Disclosure; Fig 18; 127pp; English.
XX XX
XX CC This invention describes a novel vascular endothelial growth factor-X
XX CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
XX CC antidiabetic activity and acts as an angiogenesis and vascularization
XX CC regulator. An antisense molecule of the invention is useful for treating
XX CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX CC retinopathy by inhibiting angiogenic activity or inappropriate
XX CC vascularization including formation and proliferation of new blood
XX CC vessels, growth and development of tissues, tissue regeneration and organ
XX CC and tissue repair in a subject. The products of the invention are useful
XX CC for preparing medicaments for treating wounds such as dermal ulcers,
XX CC pressure sores, venous sores, diabetic ulcers and burns and to promote
XX CC skin graft growth, tissue repair, proliferation of new blood vessels,
XX CC tissue regeneration and organ repair by promoting angiogenic activity or
XX CC vascularization. This sequence represents a human VEGF-X protein
XX CC described in the method of the invention.
XX SQ Sequence 227 AA;

Query Match 15.1%; Score 52; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 LYSCTPRNFVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCQCV 298
Db 15 lysctprnfsvsireelkrttdtiffpgcllvkrcgncacclhncncqcvp 66

RESULT 10
AAY84558
ID AAY84558 standard; Protein; 318 AA.
XX XX
XX AC AAY84558;
XX XX
XX DT 25-JUL-2000 (first entry)
XX DE A fragment of platelet-derived growth factor C (PDGF-C).
XX XX
XX KW Platelet-derived growth factor C; PDGF-C; cell proliferation;
XX KW growth factor; heparin; connective tissue; wound healing; VEGF-F;
XX KW fibroblast mitogenesis; PDGF alpha receptor activation; tumour growth;
XX KW choriocarcinoma; Wilms tumour; megakaryoblastic leukaemia;
XX KW lung carcinoma; erythroleukemia; tissue remodelling.
XX XX
XX OS Homo sapiens.
XX XX

```

DE	XX	Lung cancer associated polypeptide sequence SEQ ID 776.
KW	XX	Human; lung cancer associated protein; neuroprotective; cytostatic;
KW	XX	cardioactive; immunomodulatory; muscular active; vulnery;
KW	XX	gastrointestinal; nephrotropic; antiinfective; gynecological;
KW	XX	antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW	XX	proliferative disorder; wound healing; infectious disease.
XX	XX	
XX	OS	Homo sapiens.
XX	PN	W0200055180-A2.
XX	XX	
PD	XX	21-SEP-2000.
PF	XX	
PF	XX	08-MAR-2000; 2000WO-US05918.
XX	XX	
XX	XX	12-MAR-1999; 99US-0124270.
XX	XX	
PA	XX	(HUMA-) HUMAN GENOME SCI INC.
PA	XX	(ROSE/) ROSEN C A.
XX	XX	
PI	XX	Ruben SM;
XX	XX	
XX	XX	WPI; 2000-587514/55.
DR	XX	N-ESDB; AAF18314.
XX	XX	
PT	XX	Lung cancer associated gene sequences, referred to as lung cancer
PT	XX	antigens, useful for treatment, prevention, and diagnosis of disorders
PT	XX	such as lung cancer -
XX	XX	
PS	XX	Claim 11; Page 1305-1306; 1425pp; English.
XX	XX	
CC	XX	Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC	XX	associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC	XX	associated proteins and polynucleotide sequences, their agonists, and
CC	XX	antagonists may have neuroprotective; cytostatic; cardioactive;
CC	XX	immunomodulatory; muscular active general; vulnery; gastrointestinal
CC	XX	general; nephrotropic; antiinfective; gynecological; or antibacterial
CC	XX	activity. The invention also includes antibodies specific for the
CC	XX	protein or polynucleotide sequences. The lung cancer associated
CC	XX	polynucleotide sequences may be used for detection of lung cancer,
CC	XX	chromosome identification, as chromosome markers, and for numerous other
CC	XX	diagnostic or research purposes. The proteins may be used to treat
CC	XX	disorders such as neural, immune, muscular, reproductive,
CC	XX	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC	XX	disorders. The proteins may also be used in the treatment of wounds and
CC	XX	infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC	XX	peptide AAB58549 are used in the course of the invention for the
CC	XX	identification and characterisation of the polynucleotide and protein
XX	XX	sequences.
XX	XX	
XX	XX	Sequence 339 AA;
XX	XX	
XX	XX	Query Match 15.1%; Score 52; DB 21; Length 339;
XX	XX	Best Local Similarity 100.0%; Pred. No. 1.5e-41;
XX	XX	Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	247	LYSCTPRNFSVSIRELKRTDIFWPGCLLVKRCGNCACCLHNCNECQVP 298
Db	241	lyscptrfnfsvsiireelkrtidfpgclllvkrngncacclhncnecqvp 292
RESULT	12	
AA	AY33679	
ID	AY33679	standard; Protein; 345 AA.
XX	XX	
AC	XX	AY33679;
XX	XX	
DT	11-JAN-2000	(first entry)
XX	XX	
DE	XX	Human VEGF-E protein.
XX	XX	

KW VEGF-E; human; vascular endothelial cell growth factor; wound repair;
 KW treatment; cardiovascular disorder; endothelial disorder; therapy;
 KW tissue generation; regeneration; cardiac hypertrophy; cancer; detection;
 KW angiogenic disorder; age-related macular degeneration; vascular disease;
 KW neovascularization; tumor; gene mapping.
 XX Homo sapiens.
 OS
 XX
 XX WO9947677-A2.
 PN
 XX
 XX 23-SEP-1999.
 PD
 XX
 XX 10-MAR-1999; 99WO-US05190.
 PF
 XX
 XX 17-MAR-1998; 98US-0040220.
 PR
 XX 02-NOV-1998; 98US-0184216.
 PR
 XX (GETH) GENENTECH INC.
 PA
 XX
 XX Ferrara N, Kuo SS;
 PI
 XX
 XX WPI; 1999-580306/49.
 DR
 XX N-PSDB; AA23691.
 DR
 XX
 XX New growth factor polypeptide useful for treating cardiovascular or
 PT endothelial disorders, e.g. cardiac hypertrophy
 PT
 XX
 XX Claim 1; Fig 2; 122pp; English.
 PS
 XX
 XX This invention describes the isolation of a novel human vascular
 CC endothelial cell growth factor-E (VEGF-E) polypeptide which has
 CC tranquilizer, vulnery and cardiant activity. VEGF-E can be administered
 CC therapeutically, especially by expressing encoding polynucleotides to
 CC treat cardiovascular or endothelial disorders in mammals, especially
 CC humans. It is useful in wound repair and tissue generation and
 CC regeneration, and may especially be used to treat cardiac hypertrophy
 CC It can be combined with a carrier in pharmaceutical compositions, which
 CC can be administered to treat disorders as above. VEGF-E can be used to
 CC screen for antagonists and agonists, and the antagonists administered to
 CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
 CC age-related macular degeneration. It can be used to generate antibodies,
 CC useful therapeutically as antagonists, as above. The antibodies are also
 CC useful to detect VEGF-E polypeptide, especially to diagnose
 CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
 CC vascular disease, or neovascularization associated with tumor formation),
 CC by contacting the antibody with a tissue sample and detecting formation
 CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
 CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
 CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
 CC tissue samples. They can also be used to diagnose a disease or
 CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
 CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
 CC detecting a mutation in the VEGF-E-encoding sequence isolated from a
 CC sample. They may also be used to produce probes useful to detect related
 CC sequences or for gene mapping. This sequence represents the human VEGF-E
 CC protein described in the method of the invention.
 XX
 SQ Sequence 345 AA;

Query Match 15.1%; Score 52; DB 20; Length 345;
 Best Local Similarity 100.0%; Pred. No. 1.5e-41;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 LYSCPTPNFVSIVREELKRTDTIFWPGCLLVKRCGGNACCLHNCNECQV 298
 |||||
 Db 247 lysctprnfsivireelkrttdtlfwpgcllvkrcggncacclhncnecqcvp 298

RESULT 13
 AA41766
 ID AA41766 standard; Protein; 345 AA.
 XX

AC AAY41766;
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO200 protein sequence.
 XX
 KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KW secreted protein; transmembrane protein.
 XX Homo sapiens.
 OS
 XX
 XX WO9946281-A2.
 PN
 XX
 XX 16-SEP-1999.
 PD
 XX
 XX 08-MAR-1999; 99WO-US05028.
 PF
 XX
 XX 10-MAR-1998; 98US-0077450.
 PR
 XX 11-MAR-1998; 98US-0077632.
 PR
 XX 11-MAR-1998; 98US-0077641.
 PR
 XX 11-MAR-1998; 98US-0077649.
 PR
 XX 12-MAR-1998; 98US-0077791.
 PR
 XX 13-MAR-1998; 98US-0078004.
 PR
 XX 17-MAR-1998; 98US-0040220.
 PR
 XX 20-MAR-1998; 98US-0078886.
 PR
 XX 20-MAR-1998; 98US-0078910.
 PR
 XX 20-MAR-1998; 98US-0078936.
 PR
 XX 20-MAR-1998; 98US-0078939.
 PR
 XX 25-MAR-1998; 98US-0079294.
 PR
 XX 26-MAR-1998; 98US-0079656.
 PR
 XX 27-MAR-1998; 98US-0079663.
 PR
 XX 27-MAR-1998; 98US-0079664.
 PR
 XX 27-MAR-1998; 98US-0079689.
 PR
 XX 27-MAR-1998; 98US-0079728.
 PR
 XX 27-MAR-1998; 98US-0079786.
 PR
 XX 30-MAR-1998; 98US-0079920.
 PR
 XX 30-MAR-1998; 98US-0079923.
 PR
 XX 31-MAR-1998; 98US-0080105.
 PR
 XX 31-MAR-1998; 98US-0080107.
 PR
 XX 31-MAR-1998; 98US-0080169.
 PR
 XX 31-MAR-1998; 98US-0080194.
 PR
 XX 01-APR-1998; 98US-0080327.
 PR
 XX 01-APR-1998; 98US-0080328.
 PR
 XX 01-APR-1998; 98US-0080333.
 PR
 XX 01-APR-1998; 98US-0080334.
 PR
 XX 01-APR-1998; 98US-0081049.
 PR
 XX 08-APR-1998; 98US-0081070.
 PR
 XX 08-APR-1998; 98US-0081071.
 PR
 XX 09-APR-1998; 98US-0081195.
 PR
 XX 09-APR-1998; 98US-0081203.
 PR
 XX 09-APR-1998; 98US-0081229.
 PR
 XX 15-APR-1998; 98US-0081817.
 PR
 XX 15-APR-1998; 98US-0081838.
 PR
 XX 15-APR-1998; 98US-0081952.
 PR
 XX 15-APR-1998; 98US-0081955.
 PR
 XX 21-APR-1998; 98US-0082568.
 PR
 XX 21-APR-1998; 98US-0082569.
 PR
 XX 22-APR-1998; 98US-0082700.
 PR
 XX 22-APR-1998; 98US-0082704.
 PR
 XX 22-APR-1998; 98US-0082804.
 PR
 XX 23-APR-1998; 98US-0082767.
 PR
 XX 23-APR-1998; 98US-0082796.
 PR
 XX 27-APR-1998; 98US-0083336.
 PR
 XX 28-APR-1998; 98US-0083322.
 PR
 XX 29-APR-1998; 98US-0083392.
 PR
 XX 29-APR-1998; 98US-0083495.
 PR
 XX 29-APR-1998; 98US-0083496.
 PR
 XX 29-APR-1998; 98US-0083499.
 PR
 XX 29-APR-1998; 98US-0083500.
 PR
 XX 29-APR-1998; 98US-0083545.
 PR
 XX 29-APR-1998; 98US-0083554.
 PR
 XX 29-APR-1998; 98US-0083558.

ID	AAY30023 standard; Protein; 345 AA.
XX	
AC	AAY30023;
XX	
DT	11-OCT-1999 (first entry)
XX	
DE	Human vascular endothelial growth factor related protein.
XX	
KW	Vascular endothelial growth factor related protein; VEGF-R protein;
KW	tissue growth inhibition; tumour growth; cancer; tissue growth;
KW	angiogenesis; coronary artery blockage.
OS	Homo sapiens.
XX	
PV	WO9937671-A1.
PN	
XX	
PD	29-JUL-1999.
XX	
PF	26-JAN-1999; 99WO-US01574.
XX	
PR	31-AUG-1998; 98US-0098548.
PR	27-JAN-1998; 98US-0072635.
PR	05-JUN-1998; 98US-0088089.
PR	24-JUN-1998; 98US-0090544.
XX	
PA	(ELIL) LILLY & CO ELI.
PI	Dou S., Na S., Song HY;
PT	WI; 1999-458680/38.
DR	N-PSDB; AAX86352.
XX	
CC	A vascular endothelial growth factor related protein and related polynucleotide, useful for identifying antagonists and binding compounds
PP	Claim 1; Page 56-58; 62pp; English.
SS	
SC	The present sequence represents a vascular endothelial growth factor related (VEGF-R) protein. VEGF-R can be used in assays to identify compounds that bind to it or that antagonize its activity. VEGF-R antagonists (e.g. anti-VEGF-R antibodies) are useful for inhibiting tissue growth. This is useful for inhibiting tumour growth and for treating cancer. VEGF-R itself can be used to stimulate tissue growth, angiogenesis and to treat coronary artery blockage. The VEGF-R coding sequence can be used for the recombinant production of the VEGF-R protein.
SQ	Sequence 345 AA;
Query Match 15.1%; Score 52; DB 20; Length 345; Best Local Similarity 100.0%; Pred. No. 1.5e+41; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	247 LYSCPTRNFVSIRELKRTDITFPGCLLVKRCGNCAACCLHNCNQCQVP 298 247 lycscptrnfvsireelkrtddifwpgccllvkrccgncacclhnecqcvp 298
Db	
RESULT 15 AAB48657 ID AAB48657 standard; Protein; 345 AA. AC AAB48657; XX DT DT NO:2001 (first entry) XX DE Human zvegfg3, SEQ ID NO:33. XX KW Human; zvegf3; zvegf4 fusion; growth factor homologue; VEGF/PDGFR family; CW domain; PDGR-like activity; mitogenic; osteogenic; KW neovascularisation,tissue repair; proliferation; differentiation;	

KW liver damage; neuroregenerative; Alzheimer's disease; multiple sclerosis;
 KW periodontal disease; bone fracture; wound healing; vulnery; ischaemia;
 KW immunomodulation; hepatic.

XX Homo sapiens.

XX WO2000066736-A1.

XX PD 09-NOV-2000.

XX PF 03-MAY-2000; 2000WO-US40047.

XX PR 03-MAY-1999; 99US-0304216.

XX PR 10-NOV-1999; 99US-0164463.

XX PR 04-FEB-2000; 2000US-0180169.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;

XX DR WPI; 2000-687541/67.

XX DR N-PSDB; AAC81582.

XX Growth factor homologs and the nucleic acids that encode them, useful
 PT e.g. for treating liver damage, ischemia, multiple sclerosis and
 PT Alzheimer's disease

XX PS Claim 48; Page 125-126; 143pp; English.

XX The invention relates to the human growth factor homologue zvegfg
 CC (AAB48653), and nucleic acids encoding it (AAC81555). Zvegfg4 is a member
 CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
 CC growth factor) family. Zvegfg4 has a growth factor domain (AAB48654)
 CC characterised by a PDGF cysteine knot structure, and a CUB domain
 CC (AAB48655) which has a beta barrel structure. Zvegfg4 has PDGF-like
 CC activity, having mitogenic activity on fibroblasts, vascular smooth
 CC muscle cells and pericytes, and has also been shown to stimulate bone
 CC growth. The invention also relates to fusion proteins comprising human
 CC zvegfg4 or fragments thereof, particularly human zvegfg4/human zvegfg3
 CC fusions; expression constructs and host cells comprising human zvegfg4
 CC nucleic acids; the recombinant expression of human zvegfg4; an antibody
 CC which binds to human zvegfg4 or a fragment thereof; a method of activating
 CC a cell-surface PDGF receptor using a zvegfg4-derived polypeptide; a
 CC method of modulating the proliferation, differentiation, migration or
 CC metabolism of bone cells, comprising exposing bone cells to
 CC zvegfg4-derived polypeptides; and a method of detecting a genetic
 CC abnormality in the zvegfg4 gene of a patient. Zvegfg4 proteins and derived
 CC fragments may be used to stimulate tissue development or repair, or
 CC cellular differentiation or proliferation. They are particularly used for
 CC the treatment or repair of liver damage, and may also be used to
 CC modulate neurite growth (e.g. in the treatment of Alzheimer's disease or
 CC multiple sclerosis). Due to their osteogenic activity, they may be used
 CC in the treatment of periodontal disease and fractures. They may also be
 CC used to enhance expansion and mobilisation of haematopoietic stem cells
 CC and endothelial precursor stem cells, which may be useful in the
 CC treatment of ischaemia, in wound healing, and in the modulation of the
 CC immune system. The present sequence represents human zvegfg3.

XX Sequence 345 AA;

Query Match 15.1%; Score 52; DB 21; Length 345;
 Best Local Similarity 100.0%; Pred. No. 1.5e-41;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 LYSCPTPRNFSVSIREEELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCPV 298

|||||

Db 247 lysctprnfsvisreelkrttdifwpgcllvkrcgncacclhncncqcvp 298

Search completed: September 5, 2001, 10:56:31
 Job time: 113 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:54:58 ; Search time 12.33 Seconds
(without alignments)
576.129 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 345
Sequence: 1 MLLGLLLLSALAGRTGT.....DVALEHHECDVCRGNAGG 345

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2.6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	2.3	49	US-09-208-140-5	Sequence 5, Appli
2	8	2.3	49	US-09-208-140-6	Sequence 6, Appli
3	8	2.3	788	US-08-572-225-1	Sequence 1, Appli
4	8	2.3	855	US-08-938-365-3	Sequence 3, Appli
5	8	2.3	867	US-08-938-365-2	Sequence 2, Appli
6	8	2.3	954	US-08-749-169A-3	Sequence 3, Appli
7	8	2.3	954	US-09-130-032A-3	Sequence 3, Appli
8	8	2.3	3033	US-07-925-695-5	Sequence 5, Appli
9	7	2.0	20	US-08-078-311-9	Sequence 9, Appli
10	7	2.0	20	US-08-460-402-9	Sequence 9, Appli
11	7	2.0	24	US-08-957-001B-22	Sequence 22, Appli
12	7	2.0	24	US-09-496-301-22	Sequence 22, Appli
13	7	2.0	85	US-08-341-219-22	Sequence 22, Appli
14	7	2.0	85	US-08-912-314A-22	Sequence 22, Appli
15	7	2.0	116	US-08-053-131-185	Sequence 185, App
16	7	2.0	117	US-08-096-762-185	Sequence 185, App
17	7	2.0	117	US-09-042-353-48	Sequence 48, Appli
18	7	2.0	128	US-08-259-372A-14	Sequence 14, Appli
19	7	2.0	128	US-08-468-671-14	Sequence 14, Appli
20	7	2.0	128	US-08-470-139-8	Sequence 8, Appli
21	7	2.0	129	US-08-217-918-2	Sequence 2, Appli
22	7	2.0	155	5338678-1	Patent No. 5338678
23	7	2.0	158	5338678-2	Patent No. 5338678
24	7	2.0	184	US-08-078-311-4	Sequence 4, Appli
25	7	2.0	184	US-08-460-402-4	Sequence 4, Appli
26	7	2.0	203	US-08-855-825-12	Sequence 12, Appli
27	7	2.0	203	US-08-855-825-14	Sequence 14, Appli

28	7	2.0	205	2	US-08-912-227-4	Sequence 4, Appli
29	7	2.0	205	4	US-08-883-086-8	Sequence 8, Appli
30	7	2.0	234	2	US-07-690-192-2	Sequence 2, Appli
31	7	2.0	234	3	US-08-487-550-2	Sequence 2, Appli
32	7	2.0	235	3	US-08-812-586-16	Sequence 16, Appli
33	7	2.0	236	1	US-08-157-101A-5	Sequence 5, Appli
34	7	2.0	236	3	US-08-487-550-10	Sequence 10, Appli
35	7	2.0	239	3	US-08-487-550-6	Sequence 6, Appli
36	7	2.0	240	4	US-09-049-672A-11	Sequence 11, Appli
37	7	2.0	241	2	US-07-916-098A-56	Sequence 56, Appli
38	7	2.0	284	2	US-08-078-311-14	Sequence 14, Appli
39	7	2.0	284	2	US-08-078-311-24	Sequence 24, Appli
40	7	2.0	284	2	US-08-460-402-14	Sequence 14, Appli
41	7	2.0	284	2	US-08-460-402-24	Sequence 24, Appli
42	7	2.0	307	3	US-08-812-586-2	Sequence 2, Appli
43	7	2.0	312	2	US-09-014-969-17	Sequence 17, Appli
44	7	2.0	314	2	US-08-408-095-23	Sequence 23, Appli
45	7	2.0	355	3	US-09-082-270-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-208-140-5
; Sequence 5, Application US/09208140
; Patent No. 6228576
; GENERAL INFORMATION:
; APPLICANT: Del Vecchio, Alfred
; TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
; FILE REFERENCE: P50743
; CURRENT APPLICATION NUMBER: US/09/208.140
; CURRENT FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-09-208-140-5

Query Match 2.3%; Score 8; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LLLGLLLL 9
| | | | | | | |
Db 30 LLLGLLLL 37

RESULT 2
US-09-208-140-6
; Sequence 6, Application US/09208140
; Patent No. 6228576
; GENERAL INFORMATION:
; APPLICANT: Del Vecchio, Alfred
; TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
; FILE REFERENCE: P50743
; CURRENT APPLICATION NUMBER: US/09/208.140
; CURRENT FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-09-208-140-6

Query Match 2.3%; Score 8; DB 4; Length 49;


```
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLL 9
Db 30 LLLGLLLL 37

RESULT 3
US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojlma, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,225
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-572-225-1

Query Match 2.3%; Score 8; DB 1; Length 788;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLL 9
Db 9 LLLGLLLL 16

RESULT 4
US-08-938-365-3
; Sequence 3, Application US/08938365
; Patent No. 5989909
; GENERAL INFORMATION:
; APPLICANT: Yang, Pan
; TITLE OF INVENTION: HUCHORDIN AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,365
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/040001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-365-3

Query Match 2.3%; Score 8; DB 2; Length 855;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLL 9
Db 7 LLLGLLLL 14

RESULT 5
US-08-938-365-2
; Sequence 2, Application US/08938365
; Patent No. 5989909
; GENERAL INFORMATION:
; APPLICANT: Yang, Pan
; TITLE OF INVENTION: HUCHORDIN AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,365
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
```

REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/040001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 867 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-938-365-2

Query Match 2.3%; Score 8; DB 2; Length 867;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
Db 12 LLLGLLLL 19

RESULT 6
US-08-749-169A-3
Sequence 3, Application US/08749169A
Patent No. 5846770
GENERAL INFORMATION:
APPLICANT: RACIE, Lisa
APPLICANT: LAVALLIE, Edward
APPLICANT: DEROBERTIS, Edward
TITLE OF INVENTION: CHORDIN COMPOSITIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,169A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: GI 5284
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8260
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-749-169A-3

Query Match 2.3%; Score 8; DB 2; Length 954;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
Db 12 LLLGLLLL 19

Db 12 LLLGLLLL 19
RESULT 7
US-09-130-032A-3
Sequence 3, Application US/09130032A
Patent No. 5986056
GENERAL INFORMATION:
APPLICANT: Lavallie, Edward
APPLICANT: Racie, Lisa
APPLICANT: Derobertis, Edward
TITLE OF INVENTION: HUMAN CHORDIN COMPOSITIONS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/130,032A
FILING DATE: August 4, 1998
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: GI 5284-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8260
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-130-032A-3

Query Match 2.3%; Score 8; DB 2; Length 954;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
Db 12 LLLGLLLL 19

RESULT 8
US-07-925-695-5
Sequence 5, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
DETECTION SYSTEMS
TITLE OF INVENTION: DETECTION SYSTEMS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Weillacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: linear
US-07-925-695-5

Query Match 2.3%; Score 8; DB 1; Length 3033;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
| | | | |
DB 3014 LLLGLLLL 3021

RESULT 9
US-08-078-311-9
Sequence 9, Application US/08078311
Patent No. 5925750
GENERAL INFORMATION:
APPLICANT: Charon, Martine
APPLICANT: Gisselbrecht, Silvie
APPLICANT: Penciolelli, Jean-Francis
APPLICANT: Souyri, Michele
APPLICANT: Tambourin, Pierre
APPLICANT: Varlet, Paule
APPLICANT: Vigon, Isabelle
APPLICANT: Wendling, Francoise
TITLE OF INVENTION: Polypeptide of a Growth Factor Receptor
TITLE OF INVENTION: Family, Application in the Diagnosis and Treatment of
TITLE OF INVENTION: Myeloproliferative Disease
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 5925750west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/078,311
FILING DATE: 18-JUN-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR90/00762
FILING DATE: 19-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Kowalchuk, Katherine M.
REGISTRATION NUMBER: 36,848
REFERENCE/DOCKET NUMBER: 8076.84USWO
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-078-311-9

Query Match 2.0%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLGLLLL 9
| | | | |
DB 14 LLGLLLL 20

RESULT 10
US-08-460-402-9
Sequence 9, Application US/08460402
Patent No. 5989833
GENERAL INFORMATION:
APPLICANT: Charon, Martine
APPLICANT: Gisselbrecht, Silvie
APPLICANT: Penciolelli, Jean-Francis
APPLICANT: Souyri, Michele
APPLICANT: Tambourin, Pierre
APPLICANT: Varlet, Paule
APPLICANT: Vigon, Isabelle
APPLICANT: Wendling, Francoise
TITLE OF INVENTION: Polypeptide of a Growth Factor Receptor
TITLE OF INVENTION: Family, Application in the Diagnosis and Treatment of
TITLE OF INVENTION: Myeloproliferative Disease
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 90 South 7th Street, 3100 No. 5989833west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,402
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,259
FILING DATE: 20-SEP-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,311
FILING DATE: 18-JUN-1993
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR90/00762
FILING DATE: 14-OCT-1990
ATTORNEY/AGENT INFORMATION:

NAME: Randall A. Hillson
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.84US03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-402-9

Query Match 2.0%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLGLLLL 9
Db 14 LLGLLLL 20

RESULT 11
US-08-957-001B-22
; Sequence 22, Application US/08957001B
; Patent No. 6228621
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Madaio, Michael
; APPLICANT: Weiner, David B.
; TITLE OF INVENTION: IMPROVED VACCINES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6228621Iris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: windows
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,001B
FILING DATE: 23-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,592
FILING DATE: 23-OCT-1996
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-3303
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-957-001B-22

Query Match 2.0%; Score 7; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLGLLLL 9
Db 9 LLGLLLL 15

RESULT 12
US-09-496-301-22
; Sequence 22, Application US/09496301
; Patent No. 6248565
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Madaio, Michael
; APPLICANT: Weiner, David B.
; TITLE OF INVENTION: IMPROVED VACCINES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6248565Iris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: windows
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/496,301
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/957,001
FILING DATE: 23-OCT-1997
APPLICATION NUMBER: US 60/029,592
FILING DATE: 23-OCT-1996

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-3303
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-496-301-22

Query Match 2.0%; Score 7; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLGLLLL 9
Db 9 LLGLLLL 15

RESULT 13
US-08-341-219-22
; Sequence 22, Application US/08341219
; Patent No. 5643877
; GENERAL INFORMATION:
; APPLICANT: Zohar, Y.
; APPLICANT: Rivier, J.
; APPLICANT: Powell, J.
; APPLICANT: Sherwood, N.
; APPLICANT: Gothilf, Y.

;; TITLE OF INVENTION: Compounds and Methods For Controlling
;; TITLE OF INVENTION: Reproduction in Fish
;; NUMBER OF SEQUENCES: 26
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: N.Y.
;; COUNTRY: USA
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/341,219
;; FILING DATE: 05-DEC-1994
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A.
;; REGISTRATION NUMBER: 30742
;; REFERENCE/DOCKET NUMBER: 8399-003-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 85 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;;
US-08-341-219-22

Query Match 2.0%; Score 7; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLL 8
|||||
Db 8 LLLGLLL 14

RESULT 14
US-08-912-314A-22
; Sequence 22, Application US/08912314A
; Patent No. 6210927
; GENERAL INFORMATION:
; APPLICANT: Zohar, Y.
; APPLICANT: Rivier, J.
; APPLICANT: Powell, J.
; APPLICANT: Sherwood, N.
; APPLICANT: Gothelf, Y.
; TITLE OF INVENTION: Compounds and Methods For Controlling
; TITLE OF INVENTION: Reproduction in Fish
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/912,314A
;; FILING DATE: 30-JUN-1997
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/341,219
;; FILING DATE: 05-DEC-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A.
;; REGISTRATION NUMBER: 30742
;; REFERENCE/DOCKET NUMBER: 8399-003-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 85 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;;
US-08-912-314A-22

Query Match 2.0%; Score 7; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLL 8
|||||
Db 8 LLLGLLL 14

RESULT 15
US-08-053-131-185
; Sequence 185, Application US/08053131
; Patent No. 5661016
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Key, Robert M.
; TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,131
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14643-9-3

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 185:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-053-131-185

Query Match 2.0%; Score 7; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLGLLL 9
| | | | |
Db 9 LLGLLL 15

Search completed: September 5, 2001, 10:56:51
Job time: 113 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:56:08 ; Search time 12.76 Seconds
(without alignments)
926.187 Million cell updates/sec

Title: US-09-457-066-43

Perfect score: 345

Sequence: 1 MLLGLLLTSLAGORTGT.....DVALEHHEDCVCRNAGG 345

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	9	2.6	494	1	CPS1_HUMAN	P08686	homo sapien
2	8	2.3	20	1	CPBX_CAVPO	P34033	cavia porce
3	8	2.3	105	1	PLF4_RAT	P06765	rattus norv
4	8	2.3	124	1	EP4_CANFA	Q28894	canis fami
5	8	2.3	373	1	YU02_PICAN	P34735	pichia anqu
6	8	2.3	477	1	PN3_ADECC	Q65950	canine aden
7	8	2.3	707	1	BNP1_XENLA	P98070	xenopus lae
8	8	2.3	986	1	BNP1_HUMAN	P13497	homo sapien
9	8	2.3	991	1	BNP1_MOUSE	P98063	mus musculu
10	8	2.3	994	1	MERK_MOUSE	Q60805	mus musculu
11	8	2.3	994	1	MERK_RAT	P57097	rattus norv
12	8	2.3	999	1	DSG3_HUMAN	P32926	homo sapien
13	8	2.3	3033	1	POLG_HCVJ6	P26660	h genome po
14	7	2.0	85	1	GON2_HAPBU	P37044	haplochromi
15	7	2.0	85	1	GON2_SPAAU	P51925	sparus aura
16	7	2.0	117	1	KV1I_HUMAN	P01601	homo sapien
17	7	2.0	117	1	KV1J_HUMAN	P01602	homo sapien
18	7	2.0	129	1	KV1W_HUMAN	P04431	homo sapien
19	7	2.0	129	1	KV1X_HUMAN	P04432	homo sapien
20	7	2.0	158	1	FCY1_YEAST	Q12178	saccharomyc
21	7	2.0	184	1	MPL_MPLV	P40931	myeloprolif
22	7	2.0	190	1	BCT7_SHEEP	P50415	ovis aries
23	7	2.0	193	1	SAP3_MOUSE	Q60648	mus musculu
24	7	2.0	197	1	TNFB_RABIT	P10154	oryctolagus
25	7	2.0	204	1	TNFB_BOVIN	Q06600	bos taurus
26	7	2.0	204	1	TNFB_PIG	P26445	sus scrofa
27	7	2.0	204	1	TNFB_HUMAN	P01374	homo sapien
28	7	2.0	205	1	TNFB_ACIS7	Q44278	acinetobact
29	7	2.0	240	1	MYRA_METWA	O59640	methanosarc
30	7	2.0	254	1	KDSB_CHLTR	Q59320	chlamydia t
31	7	2.0	260	1	ER29_RAT	P52555	chlamydia t
32	7	2.0	261	1	Y612_SYNY3	P72581	synecocyst
33	7	2.0	307	1	YAC2_SCHPO	Q09816	schizosacch

RESULT 1

ID	CPS1_HUMAN	STANDARD;	PRT;	494 AA.
AC	P08686;	P04033; Q01204;		
DT	01-NOV-1986 (Rel. 03, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	CYTCHROME P450 XXIB (EC 1.14.99.10) (STEROID 21-HYDROXYLASE)			
DE	(P450-C21B).			
GN	CYP21A2 OR CYP21B OR CYP21.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86259742; PubMed=3487786;			
RA	White P.C., New M.I., Dupont B.;			
RT	"Structure of human steroid 21-hydroxylase genes.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5111-5115(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86206051; PubMed=3486422;			
RA	Higashi Y., Yoshioka H., Yamane M., Gotoh O., Fujii-Kuriyama Y.;			
RT	"Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: a pseudogene and a genuine gene.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 83:2841-2845(1986).			
RN	[3]			
RP	SEQUENCE OF 149-182 FROM N.A.			
RX	MEDLINE=85113228; PubMed=3871526;			
RA	Carroll M.C., Campbell R.D., Porter R.R.;			
RT	"Mapping of steroid 21-hydroxylase genes adjacent to complement component C4 genes in HLA, the major histocompatibility complex in man.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:521-525(1985).			
RN	[4]			
RN	MUTAGENESIS OF SER-268; VAL-281 AND CYS-428.			
RX	MEDLINE=91324486; PubMed=1864962;			
RA	Wu D.A., Chung B.C.;			
RT	"Mutations of P450c21 (steroid 21-hydroxylase) at Cys428, Val281, and Ser268 result in complete, partial, or no loss of enzymatic activity, respectively.;"			
RL	J. Clin. Invest. 88:519-523(1991).			
RN	[5]			
RP	REVIEW ON CAH-NC VARIANTS.			
RA	Gunn S.K., Sherman L.D., Therrell B.L., Overbach D.I.;			
RT	"Molecular genetics of 21-hydroxylase deficient late-onset adrenal hyperplasia.;"			
RL	Semin. Reprod. Endocrinol. 11:347-352(1993).			
RN	[6]			
RP	REVIEW ON CAH VARIANTS.			
RX	MEDLINE=94362708; PubMed=8081391;			
RA	White P.C., Tusie-Luna M.-T., New M.I., Speiser P.W.;			
RT	"Mutations in steroid 21-hydroxylase (CYP21).;"			
RL	Hum. Mutat. 3:373-378(1994).			

34	7	2.0	314	1	NIP2_HUMAN	Q12982	homo sapien
35	7	2.0	333	1	WNT3_HUMAN	P56703	homo sapien
36	7	2.0	343	1	WNT3_MOUSE	Q61327	mus musculu
37	7	2.0	355	1	WNT3_MOUSE	P17553	mus musculu
38	7	2.0	359	1	CD72_HUMAN	P21854	homo sapien
39	7	2.0	388	1	GYRB_ACIIHA	Q44065	acinetobact
40	7	2.0	397	1	ISP7_SCHPO	P40902	schizosacch
41	7	2.0	423	1	YCBD_ECOLI	P31545	escherichia
42	7	2.0	438	1	LCAT_MOUSE	P16301	mus musculu
43	7	2.0	440	1	LCAT_HUMAN	P04180	homo sapien
44	7	2.0	440	1	LCAT_PAPAN	Q08758	papio anubi
45	7	2.0	440	1	LCAT_RABIT	P53761	oryctolagus

ALIGNMENTS

RP SEQUENCE FROM N.A., AND VARIANTS CAH THR-268 AND SER-493.
RX MEDLINE=87275858; PubMed=3038528;
RA Rodrigues N.R., Dunham I., Yu C.Y., Carroll M.C., Porter R.R.,
RA Campbell R.D.;
RT "Molecular characterization of the HLA-linked steroid 21-hydroxylase
B gene from an individual with congenital adrenal hyperplasia.";
RL EMO J. 6:1653-1661(1987).
[8]
RP VARIANT CAH LEU-30, AND VARIANT THR-268.
RX MEDLINE=91304433; PubMed=2072928;
RA Tusie-Luna M.T., Speiser P.W., Dumic M., New M.I., White P.C.;
RT "A mutation (Pro-30 to Leu) in CYP21 represents a potential
nonclassical steroid 21-hydroxylase deficiency allele.";
RL Mol. Endocrinol. 5:685-692(1991).
[9]
RP VARIANT CAH ASN-172.
RX MEDLINE=8814483; PubMed=3257825;
RA Amor M., Parker K.L., Globerman H., New M.I., White P.C.;
RT "Mutation in the CYP21B gene (Ile-172-->Asn) causes steroid 21-
hydroxylase deficiency.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1600-1604(1988).
[10]
RP VARIANT CAH ASN-172.
RX MEDLINE=92039635; PubMed=1937474;
RA Partanen J., Campbell R.D.;
RT "Substitution of Ile-172 to Asn in the steroid 21-hydroxylase B
(P450c21b) gene in a Finnish patient with the simple virilizing form
of congenital adrenal hyperplasia.";
RL Hum. Genet. 87:716-720(1991).
[11]
RP VARIANTS CAH ASN-172 AND TRP-356.
RX MEDLINE=90154020; PubMed=2303461;
RA Chiou S.-H., Hu M.-C., Chung B.-C.;
RT "A missense mutation at Ile172-->Asn or Arg356-->Trp causes steroid
21-hydroxylase deficiency.";
RL J. Biol. Chem. 265:3549-3552(1990).
[12]
RP VARIANTS CAH-NC L-30; N-172; N-236; E-237; K-239; L-281 AND W-356.
RX MEDLINE=92355796; PubMed=1644925;
RA Speiser P.W., Dupont J., Zhu D., Serrat J., Buegeleisen M.,
RA Tusie-Luna M.-T., Lesser M., New M.I., White P.C.;
RT "Disease expression and molecular genotype in congenital adrenal
hyperplasia due to 21-hydroxylase deficiency.";
RL J. Clin. Invest. 90:584-595(1992).
[13]
RP VARIANTS CAH HIS-339 AND SER-453.
RX MEDLINE=93024490; PubMed=1406709;
RA Helmeberg A., Tusie-Luna M.-T., Tabarelli M., Kofler R., White P.C.;
RT "R339H and P453S: CYP21 mutations associated with nonclassical steroid
21-hydroxylase deficiency that are not apparent gene conversions.";
RL Mol. Endocrinol. 6:1318-1322(1992).
[14]
RP VARIANT CAH SER-453.
RX MEDLINE=93024478; PubMed=1406699;
RA Overbach D., Sherman L., Ballard A.L., Azziz R.;
RT "Pro-453 to Ser mutation in CYP21 is associated with nonclassical
steroid 21-hydroxylase deficiency.";
RL Mol. Endocrinol. 6:1211-1215(1992).
[15]
RP VARIANTS CAH LEU-105; SER-291 AND SER-453.
RX MEDLINE=92357805; PubMed=1496017;
RA Wedell A., Ritzén E.M., Haglund-Stengler B., Luthman H.;
RT "Steroid 21-hydroxylase deficiency: three additional mutated alleles
and establishment of phenotype-genotype relationships of common
mutations.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7232-7236(1992).
[16]
RP VARIANT CAH PRO-483.
RX MEDLINE=93239181; PubMed=8478006;
RA Wedell A., Luthman H.;
RT "Steroid 21-hydroxylase (P450c21): a new allele and spread of
mutations through the pseudogene.";

RL Hum. Genet. 91:236-240(1993).
[17]
RP VARIANTS CAH ASN-172; ASN-236; LEU-281; PRO-483 AND SER-493.
RX MEDLINE=95268421; PubMed=7749410;
RA Barbat B., Bogyo A., Raux-Demay M.-C., Kuttann F., Boue J.,
RA Simon-Bouy B., Serre J.-L., Boue A., Mornet E.;
RT "Screening of CYP21 gene mutations in 129 French patients affected by
steroid 21-hydroxylase deficiency.";
RL Hum. Mutat. 5:126-130(1995).
[18]
RP VARIANT CAH ASP-380.
RX MEDLINE=97220598; PubMed=9067760;
RA Kirby-Keyser L., Porter C.C., Donohoue P.A.;
RT "E380D: a novel point mutation of CYP21 in an HLA-homozygous patient
with salt-losing congenital adrenal hyperplasia due to 21-hydroxylase
deficiency.";
RL Hum. Mutat. 9:181-182(1997).
[19]
RP VARIANTS CAH PRO-356 AND GLN-356.
RX MEDLINE=97331329; PubMed=9187661;
RA Lajic S., Levo A., Nikoshkov A., Lundberg Y., Partanen J., Wedell A.;
RT "A cluster of missense mutations at Arg356 of human steroid 21-
hydroxylase may impair redox partner interaction.";
RL Hum. Genet. 99:704-709(1997).
[20]
RP VARIANTS CAH1 LEU-105 AND SER-453.
RX MEDLINE=97143247; PubMed=8989258;
RA Nikoshkov A., Lajic S., Holst M., Wedell A., Luthman H.;
RT "Synergistic effect of partially inactivating mutations in steroid
21-hydroxylase deficiency.";
RL J. Clin. Endocrinol. Metab. 82:194-199(1997).
[21]
RP VARIANTS CAH, AND VARIANTS.
RX MEDLINE=98180883; PubMed=9580109;
RA Ordóñez-Sánchez M.L., Ramírez-Jiménez S., López-Gutiérrez A.U.,
RA Riba L., Gamboa-Cardiel S., Cerrillo-Rinofosa M.,
RA Altamirano-Bustamante N., Calzada-León R., Robles-Valdes C.,
RA Mendoza-Morfin F., Tusie-Luna M.T.;
RT "Molecular genetic analysis of patients carrying steroid 21-
hydroxylase deficiency in the Mexican population: identification of
possible new mutations and high prevalence of apparent germ-line
mutations.";
RL Hum. Genet. 102:170-177(1998).
[22]
RP VARIANTS CAH GLU-196 DEL; SER-291 AND PRO-483.
RX MEDLINE=98165791; PubMed=9497336;
RA Nikoshkov A., Lajic S., Vlamis-Gardikas A., Tranebjærg L., Holst M.,
RA Wedell A., Luthman H.;
RT "Naturally occurring mutants of human steroid 21-hydroxylase (P450c21)
pinpoint residues important for enzyme activity and stability.";
RL J. Biol. Chem. 273:6163-6165(1998).
[23]
RP VARIANTS CAH TYR-169; LEU-281 AND GLN-356.
RX MEDLINE=99140269; PubMed=10094562;
RA Witel S.F., Smith R., Suda-Hartman M.;
RT "Identification of CYP21 mutations, one novel, by single strand
conformational polymorphism (SSCP) analysis.";
RL Hum. Mutat. 13:172-172(1999).
[24]
RP VARIANTS CAH L-30; E-64; N-172; N-236; L-281; S-291; W-356; V-362.
RX MEDLINE=9933263; PubMed=10408778;
RA Ohlsson G., Mueller J., Skakkebaek N.E., Schwartz M.;
RT "Steroid 21-hydroxylase deficiency: mutational spectrum in Denmark,
three novel mutations, and in vitro expression analysis.";
RL Hum. Mutat. 13:482-486(1999).
[25]
RP VARIANTS CAH L-30; N-172; N-236; E-237; K-239; L-281 AND W-356.
RX MEDLINE=99335271; PubMed=10408786;
RA Kapelari K., Ghanaati Z., Wollmann H., Ventz M., Ranke M.B.,
RA Kofler R., Peters H.;
RT "A rapid screening for steroid 21-hydroxylase mutations in patients
with congenital adrenal hyperplasia.";
RL Hum. Mutat. 13:505-505(1999).

Query Match 2.6%; Score 9; DB 1; Length 494;
 Best Local Similarity 100.0%; Pred. No. 0.44; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Qy 1 MLLGLLLLL 9
 Db 1 MLLGLLLLL 9

RESULT 2

CPBX_CAVPO ID CPBX_CAVPO STANDARD; PRT; 20 AA.
 AC P34033;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CYTOCHROME P450IIB (EC 1.14.14.1) (FRAGMENT).
 OS Cavia porcellus (Guinea pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 CC NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE, AND CHARACTERIZATION.
 RC STRAIN-HARTLEY; TISSUE=Liver;
 RA Narimatsu S., Akutsu Y., Matsunaga T., Watanabe K., Yamamoto I.,
 RA Yoshimura H.;
 RT "Purification of a cytochrome P450 isozyme belonging to a subfamily
 of P450 IIB from liver microsomes of guinea pigs.";
 RL Biochem. Biophys. Res. Commun. 172:607-613(1990).
 CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
 NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME IS ACTIVE
 UPON P NITROANISOLE, ANILINE, D-BENZPHETAMINE, DELTA(9)-
 TETRAHYDROCANNABINOL (THC) AND STRYCHNINE.
 CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
 OXIDIZED FLAVOPROTEIN + H(2)O.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR PIR: A36154; A36154.
 DR InterPro: IPR001128; -.
 DR PROSITE: PS00086; CYTOCHROME_P450; PARTIAL.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT NON_TER 20
 FT SEQUENCE 20 AA; 2259 MW; 78DC81280C970A55 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLLL 9
 Db 11 LLLGLLLLL 18

RESULT 3

PLF4_RAT ID PLF4_RAT STANDARD; PRT; 105 AA.
 AC P06765;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PLATELET FACTOR 4 PRECURSOR (PF-4).
 GN SCYB4 OR PF4.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE-87144262; PubMed=3821732;
 RX Doi T., Greenberg S.M., Rosenberg R.D.;
 RT "Structure of the rat platelet factor 4 gene: a marker for
 megakaryocyte differentiation.";
 RL Mol. Cell. Biol. 7:898-904(1987).
 RN [2]
 RP O-GLYCOSYLATION.
 RX MEDLINE-94307262; PubMed=8033893;
 RA Ravanat C., Gachet C., Herbert J.-M., Schuhler S., Guillemot J.-C.,
 RA Uzabiaga F., Picard C., Ferrera P., Freund M., Cazenave J.-P.;
 RT "Rat platelets contain glycosylated and non-glycosylated forms of
 platelet factor 4. Identification and characterization by mass
 spectrometry.";
 RL Eur. J. Biochem. 223:203-210(1994).
 CC -!- FUNCTION: PLATELET FACTOR 4, NONCOVALENTLY BOUND TO A PROTEOGLYCAN
 MOLECULE, IS RELEASED DURING PLATELET AGGREGATION. PF4 NEUTRALIZES
 THE ANTICOAGULANT EFFECT OF HEPARIN BECAUSE IT BINDS MORE STRONGLY
 TO HEPARIN THAN TO THE CHONDROITIN-4-SULFATE CHAINS OF THE CARRIER
 MOLECULE. CHEMOTACTIC FOR NEUTROPHILS AND MONOCYTES.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- PTM: O-GLYCAN CONSISTS OF GAL-GALNAC DISACCHARIDE WHICH IS
 MODIFIED WITH SIALIC ACID RESIDUES (MACROHETEROGENEITY).
 CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
 C-X-C) (CHEMOKINE CXCL).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

EMBL: M15254; AAA41832.1; -.
 DR PIR: A26774; A26774.
 DR HSP: P02776; IRHP.
 DR InterPro: IPR001089; -.
 DR InterPro: IPR001811; -.
 DR Pfam: PF00048; IL8; 1.
 DR PRINTS: PR00437; SMALLCYTKCX.
 DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
 KW Cytokine; Platelet; Proteoglycan; Heparin-binding; Chemotaxis;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 105 PLATELET FACTOR 4.
 FT DISULFID 44 71 BY SIMILARITY.
 FT DISULFID 46 87 BY SIMILARITY.
 FT CARBOHYD 31 31 O-LINKED (GALNAC...) (PARTIAL).
 SQ SEQUENCE 105 AA; 11286 MW; D9CCAD2B6A284496 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLGLLLLL 9
 Db 17 LLLGLLLLL 24

RESULT 4

EP4_CANFA ID EP4_CANFA STANDARD; PRT; 124 AA.
 AC Q28894;
 DT 13-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4 PRECURSOR (CE4) (EPIDIDYMAL
 SECRETORY PROTEIN E4).
 OS Canis familiaris (Dog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 CC NCBI_TaxID=9615;

```

RN SEQUENCE FROM N.A.
RC TISSUE=Epididymis;
RX MEDLINE=95263175; PubMed=7744511;
RA Ellerbrock K., Pera I., Hartung S., Ivell R.;
RT "Gene expression in the dog epididymis: a model for human epididymal
RL function.";
RL Int. J. Androl. 17:314-323(1994).
RN RN
RP TISSUE SPECIFICITY.
RX MEDLINE=95263176; PubMed=7744512;
RA Pera I., Ivell R., Kirchhoff C.;
RT "Regional variation of specific gene expression in the dog epididymis
RL as revealed by in-situ transcript hybridization.";
RL Int. J. Androl. 17:324-330(1994).
CC -1- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.
CC -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EPIDIDYMIS. HIGHEST LEVELS ARE FOUND
CC IN THE CAPUT AND PROXIMAL CAUDA REGIONS. LOWER LEVELS IN THE
CC DISTAL CAUDA. NOT DETECTED IN THE EFFERENT DUCTS.
CC -1- SIMILARITY: BELONGS TO THE WAP-TYPE 'FOUR-DISULFIDE CORE' FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; S77395; AAB34264.1;
DR InterPro; IPR002221;
DR Pfam; PF00095; wap; 2.
DR PRINTS; PR00003; 4DISULPHORE.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
KW Signal; Glycoprotein.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 124 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.
FT DOMAIN 27 75 WAP 1.
FT DOMAIN 76 124 WAP 2.
FT CARBOHYD 44 44 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 124 AA; 12951 MW; 15AAF315B813958C CRC64;

Query Match 2.3%; Score 8; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
DB 13 LLLGLLLL 20

RESULT 5
YL02_PICAN STANDARD; PRT; 373 AA.
AC P34735;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN IN LEU2 3'REGION (FRAGMENT).
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-5445;
RX MEDLINE=95028149; PubMed=7941737;
RA Agaphonov M.O., Poznyakovski A.I., Bogdanova A.I., Ter-Avanesyan M.D.;
RT "Isolation and characterization of the LEU2 gene of Hansenula
RL polymorpha.";
RL Yeast 10:509-513(1994).
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U00889; AAA19110.1;
DR PIR; S43455; S43455.
DR InterPro; IPR002889;
DR Pfam; PF01822; WSC; 1.
KW Hypothetical protein.
FT DOMAIN 123 244 SER/THR-RICH.
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 38575 MW; 4E955FFF5D191750 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
DB 270 LLLGLLLL 277

RESULT 6
PEN3_ADECC STANDARD; PRT; 477 AA.
AC Q65950;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PENTON PROTEIN (VIRION COMPONENT III) (PENTON BASE PROTEIN).
GN PIII.
OS Canine adenovirus type 1 (strain CLL), and
OS Canine adenovirus type 1 (strain RI261).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=69150, 69151;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CLL;
RA Campbell J.B., Zhao Y.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=RI261;
RX MEDLINE=97275900; PubMed=9129661;
RA Morrison M.D., Onions D.E., Nicolson L.;
RT "Complete DNA sequence of canine adenovirus type 1.";
RL J. Gen. Virol. 78:873-878(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U55001; AAB05438.1;
DR EMBL; Y07760; CAA69061.1;
DR InterPro; IPR002605;
DR Pfam; PF01686; Adeno_Penton_B; 1.
KW Late protein.
SQ SEQUENCE 477 AA; 53464 MW; B9FDF37407D0FDA9 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 477;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 183 LDLNNV 190
 Db 163 LDLNNV 170

RESULT 7
 BMPL_XENLA STANDARD; PRT; 707 AA.
 ID BMPL_XENLA STANDARD; PRT; 707 AA.
 AC P98070;
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.-) (BMP-1).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Eupidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94085787; PubMed=8262384;
 RA Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;
 RT "Cloning and expression of cDNA encoding Xenopus laevis bone
 morphogenetic protein-1 during early embryonic development.";
 RL Gene 134:257-261(1993).
 CC -!- FUNCTION: INVOLVED IN PATTERN FORMATION IN GASTRULA AND LATER
 CC -!- DIFFERENTIATION OF DEVELOPING ORGANS.
 CC -!- DEVELOPMENTAL STAGE: BLASTULA, EARLY GASTRULA AND HATCHED
 CC -!- TADPOLES; LITTLE OR NO EXPRESSION IN MORULA AND LATE GASTRULA.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 3 CUB DOMAINS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
 CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L12249; AAA16313.1; -
 DR HSP; P00736; IAPQ.
 DR MEROPS; M12.005; -
 DR InterPro; IPR000130; -
 DR InterPro; IPR000152; -
 DR InterPro; IPR000361; -
 DR InterPro; IPR000859; -
 DR InterPro; IPR001506; -
 DR InterPro; IPR001881; -
 DR Pfam; PF01400; Astacin; 1.
 DR Pfam; PF00431; CUB; 3.
 DR Pfam; PF00008; EGF; 1.
 DR PRINTS; PR00480; ASTACIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS01180; CUB; 3.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase; Protease;
 KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP 7 83 POTENTIAL.
 FT CHAIN 84 707 BONE MORPHOGENETIC PROTEIN 1.
 FT DOMAIN 84 284 METALLOPROTEASE.
 FT DOMAIN 285 397 CUB.
 FT DOMAIN 398 509 CUB.
 FT DOMAIN 510 551 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 551

FT DOMAIN 554 666 CUB.
 FT METAL 176 176 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 177 177 BY SIMILARITY.
 FT METAL 180 180 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 186 186 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 514 526 BY SIMILARITY.
 FT DISULFID 522 535 BY SIMILARITY.
 FT DISULFID 537 550 BY SIMILARITY.
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC988D CRC64;

Query Match 2.3%; Score 8; DB 1; Length 707;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 KYDFVEV 111
 Db 607 KYDFVEV 614

RESULT 8
 BMPL_HUMAN STANDARD; PRT; 986 AA.
 ID BMPL_HUMAN STANDARD; PRT; 986 AA.
 AC P13497; Q13292; Q99421; Q99422; Q99423; Q14874;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.19) (BMP-1)
 DE (PROCOLLAGEN C-PROTEINASE) (PCP) (MAMMALIAN TOLLOID PROTEIN) (MTLD).
 GN BMP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM BMP1-3).
 RC TISSUE=Skin;
 RX PubMed=8643539;
 RA Li S.W., Sieron A.L., Fertala A., Hojima Y., Arnold W.V.,
 RA Prockop D.J.;
 RA "The C-proteinase that processes procollagens to fibrillar collagens
 RT is identical to the protein previously identified as bone morphogenic
 RT protein-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5127-5130(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM BMP1-1).
 RX MEDLINE=89072730; PubMed=3201241;
 RA Wozney J.M., Rosen V., Celeste A.J., Mitscock L.M., Whitters M.J.,
 RA Kriz R.W., Hewick R.M., Wang E.A.;
 RA "Novel regulators of bone formation: molecular clones and
 RT activities.";
 RL Science 242:1528-1534(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS BMP1-4; BMP1-5 AND BMP1-6).
 RC TISSUE=Placenta;
 RX PubMed=9500680;
 RA Janitz M., Heiser V., Boettcher U., Landt O., Lauster R.;
 RA "Three alternatively spliced variants of the gene coding for the human
 RT bone morphogenetic protein-1.";
 RL J. Mol. Med. 76:141-146(1998).
 RN [4]
 RP SEQUENCE OF 703-986 FROM N.A. (ISOFORM BMP1-3).
 RC TISSUE=Placenta;
 RX MEDLINE=95096114; PubMed=7798260;
 RA Takahara K., Lyons G.E., Greenspan D.S.;
 RA "Bone morphogenetic protein-1 and a mammalian tolloid homologue (mtld)
 RT are encoded by alternatively spliced transcripts which are
 RT differentially expressed in some tissues.";

RL J. Biol. Chem. 269:32572-32578(1994).

CC -!- FUNCTION: CLEAVES THE C-TERMINAL PROPEPTIDES OF PROCOLLAGEN I, II

CC AND II. INDUCES CARTILAGE AND BONE FORMATION.

CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF THE C-TERMINAL PROPEPTIDE AT

CC ALA-|-ASP IN TYPE I AND II PROCOLLAGENS AND AT ARG-|-ASP IN TYPE

CC III.

CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).

CC -!- ENZYME REGULATION: ACTIVITY IS INCREASED BY THE PROCOLLAGEN C-

CC ENDOPETIDASE ENHANCER PROTEIN.

CC -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; BMP1-1, BMP1-2, BMP1-3 (SHOWN

CC HERE), BMP1-4, BMP1-5 AND BMP1-6; ARE PRODUCED BY ALTERNATIVE

CC SPLICING.

CC -!- TISSUE SPECIFICITY: UBUIQUITOUS.

CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.

CC -!- SIMILARITY: CONTAINS 5 CUB DOMAINS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC

CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U50330; AAA93462.1; -

DR EMBL; M22488; AAA51833.1; -

DR EMBL; Y08723; CAA69973.1; -

DR EMBL; Y08724; CAA69974.1; -

DR EMBL; Y08725; CAA69975.1; -

DR EMBL; L35279; AAC41710.1; -

DR PIR; A37278; A37278.

DR HSPS; P00736; IAPQ.

DR MEROPS; M12.005; -

DR MIM; 112264; -

DR InterPro; IPR000130; -

DR InterPro; IPR000152; -

DR InterPro; IPR000561; -

DR InterPro; IPR000859; -

DR InterPro; IPR001506; -

DR InterPro; IPR001881; -

DR Pfam; PF01400; Astacin; 1.

DR Pfam; PF00431; CUB; 5.

DR Pfam; PF00008; EGF; 2.

DR PRINTS; PRO0480; ASTACIN.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

DR PROSITE; PS01180; CUB; 5.

DR PROSITE; PS00010; ASX_HYDROXYL; 2.

DR PROSITE; PS00022; EGF_1; FALSE_NEG.

DR PROSITE; PS01186; EGF_2; 2.

DR PROSITE; PS01187; EGF_CA; 2.

KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase;

KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;

KW Glycoprotein; Zymogen; Alternative splicing.

FT SIGNAL 1 22

FT PROPEP 23 120

FT CHAIN 121 986

FT DOMAIN 121 321

FT DOMAIN 322 434

FT DOMAIN 435 546

FT DOMAIN 547 588

FT DOMAIN 591 703

FT DOMAIN 704 743

FT DOMAIN 747 859

FT DOMAIN 860 976

FT METAL 213 213

FT ACT_SITE 214 214

FT METAL 217 217

FT METAL 223 223

FT METAL 322 348

FT DISULFID 375 397

FT DISULFID 435 461

FT DISULFID 488 510 BY SIMILARITY.

FT DISULFID 551 563 BY SIMILARITY.

FT DISULFID 559 572 BY SIMILARITY.

FT DISULFID 574 587 BY SIMILARITY.

FT DISULFID 591 617 BY SIMILARITY.

FT DISULFID 644 666 BY SIMILARITY.

FT DISULFID 707 718 BY SIMILARITY.

FT DISULFID 714 727 BY SIMILARITY.

FT DISULFID 729 742 BY SIMILARITY.

FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 599 599 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPPLIC 703 730 DKDESKDNGGCGDCQCVNTEGSEYECQR -> EKRPALQPP

FT VARSPPLIC 731 986 RGRPHOLKFRVQRNRTPQ (IN ISOFORM BMP1-1).

FT VARSPPLIC 731 986 MISSING (IN ISOFORM BMP1-1).

FT VARSPPLIC 245 302 QEYNFLKMEQVEVSELSGETYDFDSIMHYANTFSRGIFLDT

FT VARSPPLIC 303 986 IVPKYEYNGVKPIGOR -> VILHSLLLSGSRNGASFP

FT VARSPPLIC 589 622 CSLESSTHQAICWTGLFLRSPFPPLPLAAPTILRAGV

FT VARSPPLIC 623 986 (IN ISOFORM BMP1-4).

FT VARSPPLIC 703 717 AACGGFTKLNGSITSPGWKEYPPNKNCIWLQV -> GCY

FT VARSPPLIC 718 986 DLQVGKPLLDWRHCFRLSTHGPEMLGTALRG (IN

FT VARSPPLIC 748 748 MISSING (IN ISOFORM BMP1-5).

FT CONFLICT 748 748 MISSING (IN ISOFORM BMP1-6).

FT CONFLICT 934 934 D -> N (IN REF. 4).

FT CONFLICT 986 AA; 111248 MW; F89201913AC3CBEA CRC64; R -> S (IN REF. 4).

SQ SEQUENCE 986 AA; 111248 MW; F89201913AC3CBEA CRC64;

Query Match 2.3%; Score 8; DB 1; Length 986;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9

DB 9 LLLGLLLL 16

RESULT 9

BMP1_MOUSE STANDARD; PRT; 991 AA.

ID BMP1_MOUSE AC P98063;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.19) (BMP-1)

DE (PROCOLLAGEN C-PROTEINASE) (PCP) (MAMMALIAN TOLLOID PROTEIN) (MTLD).

GN BMP1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Embryo;

RC MEDLINE=94229342; PubMed=8174772;

RA Fukagawa M., Noboru S., Hogan B.L.M., Jones C.M.;

RT "Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1),

RT which is related to the Drosophila dorsoventral gene tolloid and

RT encodes a putative astacin metalloendopeptidase."

RL Dev. Biol. 163:175-183(1994).

CC -!- FUNCTION: CLEAVES THE C-TERMINAL PROPEPTIDES OF PROCOLLAGEN I, II

CC AND II. INDUCES CARTILAGE AND BONE FORMATION.

CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF THE C-TERMINAL PROPEPTIDE AT

CC ALA-|-ASP IN TYPE I AND II PROCOLLAGENS AND AT ARG-|-ASP IN TYPE

CC III.

CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).

CC -!- ENZYME REGULATION: ACTIVITY IS INCREASED BY THE PROCOLLAGEN C-

CC CC ENDOPEPTIDASE ENHANCER PROTEIN.
CC CC -1- TISSUE SPECIFICITY: AT HIGH LEVELS IN EMBRYONIC MATERNAL DECIDUUM
CC CC AND FLOOR PLATE REGION OF THE NEURAL TUBE. LESS IN DEVELOPING
CC CC MEMBRANOUS AND ENDOCHONDRAL BONE, SUBMUCOSA OF INTESTINE, DERMIS
CC CC OF SKIN AND THE MESENCHYME OF SPLEEN AND LUNG.
CC CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC CC -1- SIMILARITY: CONTAINS 5 CUB DOMAINS.
CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
CC CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: L24755; AAA37306.1; -
CC CC HSP; P00736; IAPQ.
CC CC MEROPS: M12.005; -
CC CC MGD; MGI:88176; Empl.
CC CC InterPro: IPR000130; -
CC CC InterPro: IPR000152; -
CC CC InterPro: IPR000561; -
CC CC InterPro: IPR000859; -
CC CC InterPro: IPR001506; -
CC CC InterPro: IPR001881; -
CC CC Pfam: PF01400; Astacin; 1.
CC CC Pfam: PF00431; CUB; 5.
CC CC Pfam: PF00008; EGF; 2.
CC CC PRINTS: PR00480; ASTACIN.
CC CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC CC PROSITE: PS01180; CUB; 5.
CC CC PROSITE: PS00010; ASX_HYDROXYL; 2.
CC CC PROSITE: PS00022; EGF_1; FALSE_NEG.
CC CC PROSITE: PS01186; EGF_2; 2.
CC CC PROSITE: PS01187; EGF_CA; 2.
CC CC Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase;
CC CC Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
CC CC Glycoprotein; zymogen.
CC CC SIGNAL 1 25
CC CC FT PROPEP 26 125
CC CC FT CHAIN 126 991
CC CC FT DOMAIN 126 326
CC CC FT DOMAIN 327 439
CC CC FT DOMAIN 440 551
CC CC FT DOMAIN 552 593
CC CC FT DOMAIN 596 707
CC CC FT DOMAIN 708 748
CC CC FT DOMAIN 752 864
CC CC FT DOMAIN 865 981
CC CC FT METAL 218 218
CC CC FT ACT_SITE 219 219
CC CC FT METAL 222 222
CC CC FT METAL 228 228
CC CC FT DISULFID 327 353
CC CC FT DISULFID 380 402
CC CC FT DISULFID 440 466
CC CC FT DISULFID 493 515
CC CC FT DISULFID 536 568
CC CC FT DISULFID 564 577
CC CC FT DISULFID 579 592
CC CC FT DISULFID 596 622
CC CC FT DISULFID 649 671
CC CC FT DISULFID 712 723
CC CC FT DISULFID 719 732
CC CC FT DISULFID 734 747
CC CC FT CARBOHYD 96 96
CC CC FT CARBOHYD 147 147
CC CC FT CARBOHYD 337 337
CC CC FT CARBOHYD 368 368
CC CC FT CARBOHYD 604 604

SQ SEQUENCE 991 AA; 111607 MW; 68A1847783A0BB9E CRC64;

Query Match 2.3%; Score 8; DB 1; Length 991;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 CKYDFVEV 111
DB 649 CKYDFVEV 656
IIIIIIII

RESULT 10
MERK_MOUSE STANDARD; PRT; 994 AA.
AC Q60805; Q62194;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE MER PRECURSOR (EC 2.7.1.112)
DE (C-MER) (RECEPTOR TYROSINE KINASE MERTK).
GN MERTK OR MER.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=95303502; PubMed=7784083;
RA Graham D.K., Bowman G.W., Dawson T.L., Stanford W.L., Earp H.S.,
RA Snodgrass H.R.;
RT "Cloning and developmental expression analysis of the murine c-mer
RT tyrosine kinase";
RL Oncogene 10:2349-2359(1995).
RN [2]
RP SEQUENCE OF 472-994 FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RA Dowds C.A., Burks D.J., Saling P.M.;
RT "A cDNA encoding part of a novel putative receptor tyrosine kinase";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED PREDOMINANTLY IF NOT
CC EXCLUSIVELY IN THE MONOCYTIC LINEAGE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING MOST, IF NOT ALL, STAGES OF
CC EMBRYOLOGICAL DEVELOPMENT BEGINNING IN THE MORULA AND BLASTOCYST
CC AND PROGRESSING THROUGH THE YOLK SAC AND FETAL LIVER STAGES.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC AXL/UFO SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U21301; AAA80222.1; -
CC EMBL: L11625; AAA85355.1; -
CC HSP; P06213; IIRK.
CC MGD; MGI:96965; Mer.
CC InterPro: IPR000719; -
CC InterPro: IPR001245; -
CC InterPro: IPR001777; -
CC InterPro: IPR003006; -
CC Pfam: PF00041; fn3; 2.
CC Pfam: PF00047; ig; 2.
CC Pfam: PF00069; pkinase; 1.

```

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Signal; Repeat;
KW Immunoglobulin domain; Proto-oncogene.
FT SIGNAL 1 18
FT CHAIN 19 994
FT PROTO-ONCOGENE TYROSINE-PROTEIN KINASE
FT MER.
FT DOMAIN 19 497
FT TRANSMEM 498 518
FT DOMAIN 519 994
FT DOMAIN 102 177
FT DOMAIN 206 264
FT DOMAIN 279 363
FT DOMAIN 378 468
FT DOMAIN 582 852
FT NP_BIND 588 596
FT BINDING 610 610
FT ACT_SITE 718 718
FT DISULFID 109 170
FT DISULFID 213 257
FT CARBOHYD 91 91
FT CARBOHYD 108 108
FT CARBOHYD 165 165
FT CARBOHYD 202 202
FT CARBOHYD 210 210
FT CARBOHYD 229 229
FT CARBOHYD 289 289
FT CARBOHYD 311 311
FT CARBOHYD 324 324
FT CARBOHYD 331 331
FT CARBOHYD 349 349
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 437 437
FT CARBOHYD 449 449
FT MOD_RES 749 749
FT CONFLICT 473 476
FT CONFLICT 516 516
SQ SEQUENCE 994 AA; 110156 MW; 603C09FA11F76FE0 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 994;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
Db 6 LLLGLLLL 13

RESULT 11
MERK_RAT
ID MERK_RAT STANDARD; PRT; 994 AA.
AC P57037;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE MER PRECURSOR (EC 2.7.1.112)
DE (C-MER) (RECEPTOR TYROSINE KINASE MERTK).
GN MERTK OR MER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCS;
RX MEDLINE=20164303; PubMed=10699188;
RA D'Cruz P.M., Yasumura D., Weir J., Matthes M.T., Abderrahim H.,
RA Lavail M.M., Vollrath D.;
RT "Mutation of the receptor tyrosine kinase gene MERTK in the retinal
dystrophic RCS rat.";
Hum. Mol. Genet. 9:645-651(2000).
-!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
TYROSINE PHOSPHATE.
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
-!- DISEASE: DEFECTS IN MERTK ARE THE CAUSE OF RETINAL DYSTROPHY
(RDY).
-!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
AXL/UFO SUBFAMILY
-!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AF208235; AAF44060.1; -
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Signal; Repeat;
KW Immunoglobulin domain; Proto-oncogene.
FT SIGNAL 1 18
FT CHAIN 19 994
FT PROTO-ONCOGENE TYROSINE-PROTEIN KINASE
FT MER.
FT DOMAIN 19 497
FT TRANSMEM 498 518
FT DOMAIN 519 994
FT DOMAIN 102 177
FT DOMAIN 206 264
FT DOMAIN 279 363
FT DOMAIN 378 468
FT DOMAIN 582 852
FT NP_BIND 588 596
FT BINDING 610 610
FT ACT_SITE 718 718
FT DISULFID 109 170
FT DISULFID 213 257
FT CARBOHYD 91 91
FT CARBOHYD 108 108
FT CARBOHYD 165 165
FT CARBOHYD 202 202
FT CARBOHYD 210 210
FT CARBOHYD 229 229
FT CARBOHYD 289 289
FT CARBOHYD 311 311
FT CARBOHYD 324 324
FT CARBOHYD 331 331
FT CARBOHYD 349 349
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 437 437
FT CARBOHYD 449 449
FT MOD_RES 749 749
FT CONFLICT 473 476
FT CONFLICT 516 516
SQ SEQUENCE 994 AA; 110156 MW; 603C09FA11F76FE0 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 994;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 9
Db 6 LLLGLLLL 13

RESULT 12
DSG3_HUMAN
ID DSG3_HUMAN STANDARD; PRT; 999 AA.

```

AC P32926;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DESMOGLEIN 3 PRECURSOR (130 KDA PEMPHIGUS VULGARIS ANTIGEN) (PVA).
DSG3.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92069753; PubMed=1720352;
RX Anagait M., Klaus-Kovtun V., Stanley J.R.;
RT "Autoantibodies against a novel epithelial cadherin in pemphigus
RT vulgaris, a disease of cell adhesion.";
RL Cell 67:869-877(1991).
CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOSOME JUNCTIONS.
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE
CC FILAMENTS MEDIATING CELL-CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EPIDERMIS, TONGUE, TONSIL, ESOPHAGUS AND
CC CARCINOMAS.
CC -1- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS
CC (POTENTIAL).
CC -1- DISEASE: PEMPHIGUS VULGARIS (PV) IS A POTENTIALLY LETHAL SKIN
CC DISEASE IN WHICH EPIDERMAL BLISTERS OCCUR AS THE RESULT OF THE
CC LOSS OF CELL-CELL ADHESION CAUSED BY THE ACTION OF AUTOANTIBODIES
CC AGAINST DSG3.
CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY. DESMOSOMAL SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M76482; AAA60230.1; -;
DR PIR; A41088; IJHUG3.
DR HSP; P09803; 1EDH.
DR MM; 169615; -;
DR InterPro; IPR002126; -;
DR Pfam; PF00028; cadherin; 4.
DR PROSITE; PS00232; CADHERIN.1; 3.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 49 POTENTIAL.
FT CHAIN 50 999 DESMOGLEIN 3.
FT DOMAIN 50 615 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 616 640 POTENTIAL.
FT DOMAIN 641 999 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 50 158 CADHERIN 1.
FT DOMAIN 159 268 CADHERIN 2.
FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 386 499 CADHERIN 4.
FT REPEAT 910 935 DESMOGLEIN REPEAT 1.
FT REPEAT 936 966 DESMOGLEIN REPEAT 2.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 999;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 LLLGLLLL 9
| | | | | | | |

Db 624 LLLGLLLL 631
RESULT 13
POLG_HCVJ6
ID POLG_HCVJ6 STANDARD; PRT; 3033 AA.
AC P26660;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN) (P22);
DE ENVELOPE GLYCOPROTEIN E1 (GP32); ENVELOPE GLYCOPROTEIN E2
DE (GP68) (GP70) (NS1); PROTEIN P7; NONSTRUCTURAL PROTEIN NS2 (P21)
DE (EC 3.4.22.-); PROTEASE/HELICASE NS3 (P70) (EC 3.4.21.-);
DE NONSTRUCTURAL PROTEIN NS4 (P4); NONSTRUCTURAL PROTEIN NS4B (P27);
DE NONSTRUCTURAL PROTEIN NS5A (P56); NONSTRUCTURAL PROTEIN NS5B (P66)
DE (P70) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate HC-J6) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92044440; PubMed=1658196;
RA Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
RA Machida A., Miyakawa Y., Mayumi M.;
RT "Nucleotide sequence of the genomic RNA of hepatitis C virus isolated
RT from a human carrier: comparison with reported isolates for conserved
RT and divergent regions.";
RL J. Gen. Virol. 72:2697-2704(1991).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00944; BAA00792.1; -;
DR PIR; JQ1303; JQ1303.
DR HSP; P27958; 1HEI.
DR MEROPS; S29.001; -;
DR MEROPS; U39.001; -;
DR InterPro; IPR000745; -;
DR InterPro; IPR001490; -;
DR InterPro; IPR002166; -;
DR InterPro; IPR002518; -;
DR InterPro; IPR002519; -;
DR InterPro; IPR002521; -;
DR InterPro; IPR002522; -;
DR InterPro; IPR002531; -;
DR InterPro; IPR002868; -;
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF00998; HCV_RdRP; 1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_en; 1.
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
KW Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

FT INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
 FT CHAIN 1 115 CELLULAR AMINOPEPTIDASE.
 FT CHAIN 116 191 CAPSID PROTEIN C (POTENTIAL).
 FT CHAIN 192 383 MATRIX PROTEIN (POTENTIAL).
 FT CHAIN 384 733 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
 FT CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
 FT CHAIN 1011 1619 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
 FT CHAIN 1620 1866 PROTEASE/HELICASE NS3 (POTENTIAL).
 FT CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
 FT CHAIN 2018 3033 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
 FT TRANSMEM 347 369 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
 FT ACT_SITE 1087 1087 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 1111 1111 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 1169 1169 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT NP_BIND 1234 1241 ATP (POTENTIAL).
 FT SITE 1320 1323 DECH BOX.
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 558 558 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1091 1091 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2038 2038 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2811 2811 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 3033 AA; 329165 MW; F957F5CIA273BE9E CRC64;

Query Match 2.3%; Score 8; DB 1; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 22; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 2 LLLGLLLL 9
 |||||
 Db 3014 LLLGLLLL 3021

RESULT 14
 GON2_HAPBU
 ID GON2_HAPBU STANDARD; PRT; 85 AA.
 AC P37044; P20408;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GONADOLIBERIN II PRECURSOR (GONADOTROPIN-RELEASING HORMONE II)
 DE (GNRH-II) (LH-RH II) (LULIBERIN II).
 GN GNRH2.
 OS Haplochromis burtoni.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;
 OC Cichlidae; Astotitlapia.
 OX NCBI_TaxID=8153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94151343; PubMed=8108425;
 RA White S.A., Bond C.T., Francis R.C., Kasten T.L., Fernald R.D.,
 RA Adelman J.P.;
 RT "A second gene for gonadotropin-releasing hormone: cDNA and
 expression pattern in the brain."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1423-1427(1994).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ONLY ONE CELL GROUP IN THE
 CC MESENCEPHALON.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

CC EMBL; L27435; AAA74993.1; -
 CC InterPro; IPR002012; -
 CC Pfam; PF00446; GNRH; 1.
 CC PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 85 PRONADOLIBERIN II.
 FT PEPTIDE 24 33 GONADOLIBERIN II.
 FT PEPTIDE 37 64 GNRH-ASSOCIATED PEPTIDE II-1 (POTENTIAL).
 FT PEPTIDE 67 85 GNRH-ASSOCIATED PEPTIDE II-2 (POTENTIAL).
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 85 AA; 9631 MW; CF8C0EDBF277365F CRC64;

Query Match 2.0%; Score 7; DB 1; Length 85;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLL 8
 |||||
 Db 8 LLLGLLLL 14

RESULT 15
 GON2_SPAU
 ID GON2_SPAU STANDARD; PRT; 85 AA.
 AC P51925;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GONADOLIBERIN II PRECURSOR (GONADOTROPIN-RELEASING HORMONE II)
 DE (GNRH-II) (LH-RH II) (LULIBERIN II).
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidel;
 OC Sparidae; Sparus.
 OX NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99061849; PubMed=9843645;
 RA Holland M.C.H., Gothilf Y., Meiri I., King J.A., Okuzawa K.,
 RA Elizur A., Zohar Y.;
 RT "Levels of the native forms of GNRH in the pituitary of the gilthead
 RT seabream, Sparus aurata, at several characteristic stages of the
 RT gonadal cycle."
 RL Gen. Comp. Endocrinol. 112:394-405(1998).
 RN [2]
 RP SEQUENCE OF 24-33.
 RC TISSUE=Brain;
 RX MEDLINE=95083645; PubMed=7991588;
 RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
 RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
 RT "Three forms of gonadotropin-releasing hormone characterized from
 RT brains of one species."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- MASS SPECTROMETRY: MW=1236.6; METHOD=MALDI; RANGE=24-33.

CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30325; AAA75447.1; -;
DR InterPro; IPR002012; -;
DR Pfam; PF00446; Gnrh; 1.
DR PROSITE; PS00473; Gnrh; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 85 PROGNADOLIBERIN II.
FT PEPTIDE 24 33 GONADOLIBERIN II.
FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE II.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 85 AA; 9543 MW; B53165C122722CC CRC64;

Query Match 2.0%; Score 7; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLL 8
Db 8 LLLGLLL 14

Search completed: September 5, 2001, 10:58:05
Job time: 117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2001, 10:55:48 ; Search time 24.73 Seconds
(without alignments)
1845.743 Million cell updates/sec

Title: US-09-457-066-43
Perfect score: 345
Sequence: 1 MLLGLLLTSLALAGORTGT.....DVALEHHEECDCVCRGNAGG 345

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL16:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organellae:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_unclassified:*
 - 13: sp_vertebrate:*
 - 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	100.0	345	11 Q9QY71	Q9qy71 mus musculus
2	102	29.6	345	11 Q9JHV8	Q9jhw8 mus musculus
3	79	22.9	345	11 Q9EQX6	Q9eqx6 rattus norv
4	52	15.1	345	4 Q9UL22	Q9ul22 homo sapien
5	52	15.1	345	4 Q9NR11	Q9nr11 homo sapien
6	32	9.3	345	13 Q91946	Q91946 gallus gall
7	9	2.6	97	4 Q16755	Q16755 homo sapien
8	9	2.6	364	14 Q68809	Q68809 hepatitis c
9	9	2.6	364	14 Q81599	Q81599 hepatitis c
10	9	2.6	364	14 Q96851	Q96851 hepatitis c
11	9	2.6	364	14 Q96857	Q96857 hepatitis c
12	9	2.6	364	14 Q96858	Q96858 hepatitis c
13	9	2.6	364	14 Q87755	Q87755 hepatitis c
14	9	2.6	364	14 Q87789	Q87789 hepatitis c
15	9	2.6	364	14 Q68795	Q68795 hepatitis c
16	9	2.6	364	14 Q68800	Q68800 hepatitis c
17	9	2.6	364	14 Q68803	Q68803 hepatitis c
18	9	2.6	364	14 Q68812	Q68812 hepatitis c
19	9	2.6	364	14 Q68814	Q68814 hepatitis c

20	9	2.6	364	14 Q68816	Q68816 hepatitis c
21	9	2.6	371	4 Q16748	Q16748 homo sapien
22	9	2.6	371	4 Q16746	Q16746 homo sapien
23	9	2.6	495	4 Q16749	Q16749 homo sapien
24	9	2.6	495	4 Q16742	Q16742 homo sapien
25	9	2.6	495	4 Q16741	Q16741 homo sapien
26	9	2.6	495	4 Q16874	Q16874 h cytochrom
27	9	2.6	3018	14 Q39927	Q39927 hepatitis c
28	9	2.6	3019	14 Q92529	Q92529 hepatitis c
29	9	2.6	3019	14 Q68801	Q68801 hepatitis c
30	8	2.3	31	11 Q9QVA5	Q9qva5 cavia (guin
31	8	2.3	33	11 Q9QVJ4	Q9qvj4 cavia (guin
32	8	2.3	41	11 Q9QVY1	Q9qvyl cavia (guin
33	8	2.3	86	4 Q9H2W9	Q9h2w9 homo sapien
34	8	2.3	94	4 Q9H2W8	Q9h2w8 homo sapien
35	8	2.3	116	10 Q93284	Q93284 arabidopsis
36	8	2.3	221	2 Q9KU56	Q9ku56 vibrio chol
37	8	2.3	241	11 Q92135	Q92135 rattus norv
38	8	2.3	269	10 Q93641	Q93641 zea mays (m
39	8	2.3	290	10 Q9XH03	Q9xn03 arabidopsis
40	8	2.3	347	8 Q9MGJ0	Q9mgj0 darevskia b
41	8	2.3	350	4 Q9H2D3	Q9h2d3 homo sapien
42	8	2.3	364	14 Q92533	Q92533 hepatitis c
43	8	2.3	364	14 Q68792	Q68792 hepatitis c
44	8	2.3	364	14 Q68807	Q68807 hepatitis c
45	8	2.3	364	14 Q68814	Q68814 hepatitis c

ALIGNMENTS

RESULT 1
Q9QY71 ID Q9QY71 PRELIMINARY; PRT; 345 AA.
AC Q9QY71;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FALLOTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=OVARY;
RA Tsai Y.-J., Lee R.K.-K., Chen Y.-H., Lin S.-P., Cheng W.T.-K.;
RT "CDNA cloning of fallotein from mouse ovary."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117608; AAF22516.1;
DR InterPro; IPR000072;
DR InterPro; IPR000859;
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38741 MW; 3A58A1F701B84EA2 CRC64;

Query Match 100.0%; Score 345; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLLGLLLTSLALAGORTGTRAESNLSSKQLQSSDKKEQNCVQDPRHVVVTISGNSIHS	60
DB	1	MLLGLLLTSLALAGORTGTRAESNLSSKQLQSSDKKEQNCVQDPRHVVVTISGNSIHS	60
QY	61	PKPFPYPRNMVLVWRLVAVDENVRQLTFDERFGLDEDDICKYDFVEVEPSPDGSVL	120
DB	61	PKPFPYPRNMVLVWRLVAVDENVRQLTFDERFGLDEDDICKYDFVEVEPSPDGSVL	120
QY	121	GRWCSGTVPKGKTSKGNHIRFVSDVEFPSPGFCIHYSIIMPQVTTETSPSVLPSS	180

```
Db 121 GRWCGSGTVPCKQTSKGNHIRFVSDYFSPSEPGFCIHYSIIMPQVTTTSPSVLPSS 180
Qy 181 LSLDLLNNAVTAFTLEELIRYLEPDRQWVLDLSLYKPTWOLLGKAFLYGKSKVVNLNL 240
Db 181 LSLDLLNNAVTAFTLEELIRYLEPDRQWVLDLSLYKPTWOLLGKAFLYGKSKVVNLNL 240
Qy 241 LKEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPRK 300
Db 241 LKEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPRK 300
Qy 301 VTKYHEVLQRLPKTVGKGLHKSITDVALEHHEECDCVCRGNAGG 345
Db 301 VTKYHEVLQRLPKTVGKGLHKSITDVALEHHEECDCVCRGNAGG 345

RESULT 2
Q9JHV8
ID Q9JHV8 PRELIMINARY; PRT; 345 AA.
AC Q9JHV8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
GN PDGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS-WEBSTER/NIH;
RA Ding H., Wu X., Kim I., Tam P.P.L., Koh G.Y., Nagy A.;
RT "The mouse pdgfc gene: Dynamic expression in embryonic tissues during
RT organogenesis.";
RL Mech. Dev. 0:0-0(2000).
DR EMBL; AF286725; AAF91483.1; -.
DR InterPro; IPR000072; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38886 MW; FA1486BED6D362F8 CRC64;

Query Match 29.6%; Score 102; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.9e-91;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 128 TVPGKQTSKGNHIRFVSDYFSPSEPGFCIHYSIIMPQVTTTSPSVLPSSLSLDLNL 187
Db 128 TVPGKQTSKGNHIRFVSDYFSPSEPGFCIHYSIIMPQVTTTSPSVLPSSLSLDLNL 187

Qy 188 NAVTAFTLEELIRYLEPDRQWVLDLSLYKPTWOLLGKAFLY 229
Db 188 NAVTAFTLEELIRYLEPDRQWVLDLSLYKPTWOLLGKAFLY 229

RESULT 3
Q9EQX6
ID Q9EQX6 PRELIMINARY; PRT; 345 AA.
AC Q9EQX6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN RSCDGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=WISTAR; TISSUE=KIDNEY;
RA Hamada T., Ui-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDGF-B, a Novel Growth Factor Homologous to
RT SCDGF/PDGF-C/Fallotelin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
DR EMBL; AB033830; BAB19969.1; -.
SQ SEQUENCE 345 AA; 38734 MW; F296DA6E9B765D10 CRC64;

Query Match 22.9%; Score 79; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 8.6e-69;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 VNLNLLKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQ 295
Db 236 VNLNLLKEEVKLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQ 295

Qy 296 CVPRKVTKKYHEVLQRLPK 314
Db 296 CVPRKVTKKYHEVLQRLPK 314

RESULT 4
Q9UL22
ID Q9UL22 PRELIMINARY; PRT; 345 AA.
AC Q9UL22;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SECRETORY GROWTH FACTOR-LIKE PROTEIN FALLOTEIN (SPINAL CORD-DERIVED
DE GROWTH FACTOR).
GN HSCDGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Tsai Y.J., Lee R.K.K., Lin S.P.;
RT "Fallotelin, a novel growth factor like gene identified in human
RT uterus.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=203117014; PubMed=10858496;
RA Hamada T., Ui-Tei K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDGF, is a unique
RT member of the PDGF/VEGF family.";
RL PEBB Lett. 475:97-102(2000).
DR EMBL; AF091434; AAF00049.1; -.
DR EMBL; AB033831; BAB03266.1; -.
DR InterPro; IPR000072; -.
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39029 MW; CDE9E51F40633E78 CRC64;
```

Query Match 15.1%; Score 52; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 LYSCPTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQV 298
Db 247 LYSCPTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQV 298

RESULT 5

```
O9NRAL
ID Q9NRAL1 PRELIMINARY; PRT; 345 AA.
AC Q9NRAL1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Li X., Ponten A., Aase K., Karlsson L., Abramsson A., Uutela M.,
RA Backstrom G., Hellstrom M., Bostrom H., Li H., Soriano P.,
RA Beksholtz C., Heidlin C.-H., Alitalo K., Ostman A., Eriksson U.;
RT "PDGF-C is a novel protease-activated ligand for the PDGF alpha
RT receptor.";
RL Nat. Cell Biol. 0:0-0(2000).
DR EMBL; AF244813; AAF80597.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39043 MW; 590889CEA55CC5EA CRC64;

Query Match 15.1%; Score 52; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 247 LYSCPTPRNFSVSIRELKKRTDIFWPGCLLVKRCGNCACCLHNCQCVP 298
Db 247 LYSCPTPRNFSVSIRELKKRTDIFWPGCLLVKRCGNCACCLHNCQCVP 298

RESULT 6
O91946
ID Q91946 PRELIMINARY; PRT; 345 AA.
AC Q91946;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN SCDFG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=WHITE LEHORN; TISSUE=SPINAL CORD;
RX MEDLINE=20317014; PubMed=10858496;
RA Hamada T., Ui-Tei K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDFG, is a unique
RT member of the PDGF/VBGF family.";
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AB033829; BAB03265.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38940 MW; 97ACEA992BF5128C CRC64;

Query Match 9.3%; Score 32; DB 13; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 86 IQLTFDERFGLDEDDICKYDFVEVEPSDG 117
Db 86 IQLTFDERFGLDEDDICKYDFVEVEPSDG 117

RESULT 7
Q16755
ID Q16755 PRELIMINARY; PRT; 97 AA.
AC Q16755;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE STEROID 21-MONOOXYGENASE (EC 1.14.99.10) (STEROID 21-HYDROXYLASE)
DE CYTOCHROME P450 XXIA1 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RX MEDLINE=93024490; PubMed=1406709;
RA Helmsberg A., Tusie-Luna M.T., Tabarelli M., Kofler R., White P.C.;
RT "R339H and P453S: CYP21 mutations associated with nonclassic steroid
RT 21-hydroxylase deficiency that are not apparent gene conversions.";
RL Mol. Endocrinol. 6:1318-1322(1992).
CC -|- CATALYTIC ACTIVITY: A STEROID + AH(2) + O(2) -> A 21-HYDROXYSTEROID
CC + A + H(2)O.
CC -|- COFACTOR: HEME-THIOLATE.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; X58908; CAA41711.1; -.
DR InterPro; IPR001128; -.
DR Pfam; PF00067; P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON_TER 97
SQ SEQUENCE 97 AA; 11069 MW; 84A118F6680C6737 CRC64;

Query Match 2.6%; Score 9; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLLGLLLLL 9
Db 1 MLLGLLLLL 9

RESULT 8
Q68809
ID Q68809 PRELIMINARY; PRT; 364 AA.
AC Q68809;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=JFK022;
RX MEDLINE=96226020; PubMed=8627233;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA Lesmana L.A., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
RT novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
RT (11a) genetic groups.";
RL J. Gen. Virol. 77:293-301(1996).
CC -|- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
```

```

DR EMBL; D49768; BAA08602.1; -.
DR InterPro; IPR002166; -.
DR Pfam; PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 40181 MW; 74BF8B2BD95964D3 CRC64;

Query Match
Best Local Similarity 2.6%; Score 9; DB 14; Length 364;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLLT 10
DB 345 LLLGLLLLT 353

RESULT 9
ID Q81599 PRELIMINARY; PRT; 364 AA.
AC Q81599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH580;
RX Tokita H., Okamoto H., Luengrojanakul P., Varesangthip K.,
RA Chainvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Thailand classifiable into five novel
RT genotypes in the sixth (6b), seventh (7c, 7d) and ninth (9b, 9c) major
RT genetic groups.";
RL J. Gen. Virol. 76:2329-2335(1995).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL; D37855; BAA07098.1; -.
DR InterPro; IPR002166; -.
DR Pfam; PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39996 MW; BAE791D77FDA5D8E CRC64;

Query Match
Best Local Similarity 2.6%; Score 9; DB 14; Length 364;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLLT 10
DB 345 LLLGLLLLT 353

RESULT 10
ID Q96851 PRELIMINARY; PRT; 364 AA.
AC Q96851;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN506;
RX Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,

```

```

RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VN506;
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL; D87356; BAA13339.1; -.
DR InterPro; IPR002166; -.
DR Pfam; PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39940 MW; E32B6CEAA96B7BF2 CRC64;

Query Match
Best Local Similarity 2.6%; Score 9; DB 14; Length 364;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLLT 10
DB 345 LLLGLLLLT 353

RESULT 11
ID Q96855 PRELIMINARY; PRT; 364 AA.
AC Q96855;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN538;
RX Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VN538;
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL; D87360; BAA13343.1; -.
DR InterPro; IPR002166; -.
DR Pfam; PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39966 MW; 9B3AD975864B7533 CRC64;

Query Match
Best Local Similarity 2.6%; Score 9; DB 14; Length 364;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLLT 10
DB 345 LLLGLLLLT 353

```

RESULT 12
Q96857 ID Q96857 PRELIMINARY; PRT; 364 AA.
AC Q96857;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN569;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.,
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VN569;
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.,
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL: D87362; BAA13345.1; -;
DR InterPro: IPR002166; -;
DR Pfam: PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39910 MW; 139EC77C221080E5 CRC64;
Query Match 2.6%; Score 9; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LLLGLLLLT 10
Db 345 LLLGLLLLT 353
RESULT 13
Q96858 ID Q96858 PRELIMINARY; PRT; 364 AA.
AC Q96858;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN571;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.,
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VN571;
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.,
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).

CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL: D87363; BAA13346.1; -;
DR InterPro: IPR002166; -;
DR Pfam: PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39872 MW; C631B44009FBDIC9 CRC64;
Query Match 2.6%; Score 9; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LLLGLLLLT 10
Db 345 LLLGLLLLT 353
RESULT 14
P87755 ID P87755 PRELIMINARY; PRT; 364 AA.
AC P87755;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JK093;
RX MEDLINE=96226020; PubMed=8627233;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA Lesmana L.A., Miyakawa Y., Mayumi M.,
RT "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
(11a) genetic groups.";
RL J. Gen. Virol. 77:293-301(1996).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL: D49771; BAA08605.1; -;
DR InterPro: IPR002166; -;
DR Pfam: PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 40096 MW; 141A3AF31DDB7002 CRC64;
Query Match 2.6%; Score 9; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LLLGLLLLT 10
Db 345 LLLGLLLLT 353
RESULT 15
Q68789 ID Q68789 PRELIMINARY; PRT; 364 AA.
AC Q68789;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOME POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JK012;

RX MEDLINE-96226020; PubMed-8627233;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA Lesmana L.A., Miyakawa Y., Mayumi M.,
RT "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
RT novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
RT (11a) genetic groups.",
RL J. Gen. Virol. 77:293-301(1996).
CC -!- SIMILARITY: TO HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE.
DR EMBL; D49759; BAA08593.1; -
DR InterPro: IPR002166; -
DR Pfam; PF00998; HCV_RdRp; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 40072 MW; 0D7E2B9B49023FBB CRC64;

Query Match 2.68; Score 9; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLGLLLLT 10
|||
Db 345 LLLGLLLLT 353

Search completed: September 5, 2001, 10:57:45
Job time: 117 sec